# Analysis of the Structure and Function of the *S. pombe* DNA ligase I Protein Cdc17

by

Ina Verena Martin

Thesis presented for the Degree of

Doctor of Philosophy

at the University of Edinburgh

April 2002

"The best way to have a good idea is to have lots of ideas."

Linus Pauling, 1901-1994

**For Mum and Dad**

## Acknowledgements

Firstly, I thank my supervisor Stuart MacNeill for his continuous support, advice and helpful discussions throughout this project and for critical reading of the manuscript.

Thanks are due to all members of the MacNeill lab, especially Leo Ciufo, Jess Worthington, Fiona Gray and Javier Sanchez García for their advice, assistance and helpful discussions mostly administered over a cup of tea. I would also like to thank members of the Adachi, Hardwick and Ohkura labs for helping me in various ways.

Special thanks to Joan Davidson for superb technical support.

I am indebted to the Darwin Trust of Edinburgh for financial support.
Many thanks also go to my second supervisor David Dryden for his continuous support and for encouraging me to do a PhD in the first place.

And not the least, I am thanking my family and friends for their never-ending encouragment and love.

## Abstract

DNA ligases join breaks in double-stranded DNA and are therefore crucial enzymes in all aspects of DNA metabolism. Eukaryotic DNA ligase I homologues belong to the family of ATP-dependent ligases and join Okazaki fragments generated on the lagging strand during chromosomal DNA replication. DNA ligase I enzymes consist of C-terminal catalytic domains conserved across all ATP-dependent ligases, a middle conserved domain of unknown function and N-terminal extensions, which share a conserved PCNA binding motif, but otherwise show very limited sequence similarity.

In this work structure-function analyses on the DNA ligase I homologue Cdc17 of *Schizosaccharomyces pombe* were performed. The presence of the middle conserved non-catalytic domain in addition to the catalytic domains was found to be essential for rescue of a *cdc17* deletion strain. The N-terminal domain targets the enzyme to the nucleus and mitochondria and essential residues were identified which are required for targeting to both cellular compartments. Evidence suggests that mitochondrial Cdc17 function is required for survival of *S. pombe*. The PCNA binding motif is functional *in vivo* since its absence reduces cell viability. Despite these identified functions overexpression of truncations lacking the N-terminal domain can complement a *cdc17* deletion strain.

## Table of Contents

# Abbreviations

| | |
|---|---|
| ~ | approximately |
| °C | degree Celsius |
| µ- | micro |
| µF | micro Farad |
| 5-FOA | 5-fluoroorotic acid |
| A | adenine |
| aa | amino acids |
| AMPS | ammoniumperoxodisulphate |
| AP | apurinic/apyrimidinic |
| ATP | adenosine triphosphate |
| BER | base-excision repair |
| BSA | bovine serum albumine |
| C | cytosine |
| cdc | cell division cycle |
| CDK | cyclin-dependent protein kinase |
| CNCD | conserved non-catalytic domain (of Cdc17) |
| Core | catalytic domains (of Cdc17) |
| DAPI | 4',6-diamidino-2-phenylindole dihydrochloride |
| DMF | dimethylformamide |
| DMSO | dimethylsulphoxide |
| DNA | deoxyribonucleic acid |
| EDTA | ethylen-diamine-tetraacetic acid |
| EMM | Edinburgh minimal medium |
| G | guanine |
| g | gram |
| GFP | green fluorescent protein |
| GST | glutathione-S-transferase |
| $H_2O$ | water (distilled) |
| IPTG | isopropylthio-β-D-galactoside |
| kb | kilobase |
| kDa | kilo Dalton |
| kV | kilo Volt |
| l | litre |
| LB | Luria broth |
| LiAc | lithiumacetate |
| M | molar |
| m- | milli |
| max. | maximum |
| Mb | megabase |
| MCM | mini chromosome maintenance |
| ME | malt extract |
| MPP | mitochondrial processing peptidase |
| mRNA | messenger RNA |
| mt | mitochondrial |
| MTS | mitochondrial targeting sequence |
| n- | nano |
| $NAD^+$ | nicotinic adenine dinucleotide |
| NER | nucleotide-excision repair |
| NHEJ | non-homologous end joining |
| NLS | nuclear localisation signal |
| NT | N-terminal domain (of Cdc17) |
| OD | optical density |
| ORC | origin recognition complex |
| ORF | open reading frame |

| | |
|---|---|
| *ori* | origin of replication |
| PAGE | polyacrylamide gel electrophoresis |
| PCNA | proliferating cell nuclear antigen |
| PCR | polymerase chain reaction |
| PEG | polyethylenglycol |
| PMSF | phenylmethylsulphonylfluoride |
| pol | DNA polymerase |
| PVDF | polyvinylidene difluoride |
| res | resistant |
| RF-C | replication factor C |
| RNA | ribonucleic acid |
| RPA | replication protein A |
| rpm | revolutions per minute |
| SDS | sodium dodecylsulphate |
| T | thymine |
| TEMED | N,N,N',N'-Tetramethylethylendiamine |
| tRNA | transfer RNA |
| ts | temperature-sensitive |
| UV | ultraviolet light |
| UVER | UVDE-dependent excision repair |
| v/v | volume per volume |
| w/v | weight per volume |
| X-gal | 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside |
| YE | yeast extract |
| Ω | ohm (unit of resistance) |

# 1. Introduction

## 1.1 Eukaryotic DNA replication

### 1.1.1 Overview

Chromosomal DNA replication in eukaryotes is a crucial event in the cell cycle that is responsible for the accurate and efficient duplication of the genome prior to cell division. The duplication of the genetic material occurs during S phase of the eukaryotic cell cycle and it is separated from mitosis through the gap phases G1 and G2. DNA replication involves the highly co-ordinated action of multiple replication proteins at the replication fork through a complex network of protein-protein interactions. DNA synthesis is coupled to cell cycle progression and DNA repair in order to ensure that the complete genome is replicated once and only once, thus maintaining genome integrity.

Initiation of DNA replication is regulated by the cell cycle machinery and is brought about through the dynamic assembly and disassembly of protein complexes at sites of replication initiation. Synthesis of eukaryotic DNA is achieved by numerous replication forks initiating at multiple specific sites scattered throughout each chromosome. The mode of DNA replication is semi-discontinuous: the leading strand is synthesised continuously whereas the lagging strand is formed as small DNA intermediates, termed Okazaki fragments, which are subsequently joined together. The mechanisms of leading strand synthesis and that of each Okazaki fragment are very similar: a protein-complex sometimes referred to as the primosome produces an RNA-DNA primer which is then elongated by a replicative DNA polymerase-containing protein complex. Finally, the Okazaki fragment intermediates are processed to form a continuous DNA strand (reviewed in Davey and O'Donnell, 2000).

Much of what is known about the mechanisms of DNA replication is derived from work on the model DNA replication system Simian virus 40 (SV40) using purified cellular replication proteins (reviewed in Hassell and Brinton, 1996 and Waga and Stillman, 1998). Complementing these studies, genetical and biochemical analyses of

yeast DNA replication proteins, have greatly extended the current understanding of the molecular basis of the replication process (reviewed in MacNeill and Burgers, 2000).

### 1.1.2 Initiation

DNA synthesis starts from multiple specific sites on each chromosome known as origins of replication, *ori*. Eukaryotic *ori* are bound by multi-protein origin-recognition-complexes (ORC). ORC were first described as six-subunit complexes in budding yeast (Bell and Stillman, 1992) and were shown to be present at origin DNA throughout the cell cycle (Liang and Stillman, 1997). Identification of ORC subunits in other organisms have demonstrated that ORC seem to be conserved across eukaryotes (Gavin *et al.*, 1995) and the six-subunit ORC of *S. pombe* and humans, among other species, have been purified (Moon *et al.*, 1999, Vashee *et al.*, 2001).

In *S. cerevisiae* ORC bind to Cdc6p, another essential DNA replication initiation protein (Liang *et al.*, 1995). In conjunction with the Cdt1 protein, the so-called minichromosome-maintenance (MCM) proteins are recruited to ORC (Tanaka *et al.*, 1997, Nishitani *et al.*, 2000, Maiorano *et al.*, 2000). At this stage the origin DNA becomes 'licensed' for DNA replication. Coupled to cell cycle signals, which, in *S. cerevisiae*, include the Cdc45 protein and the action of the Cdc7/Dbf4 kinase and S-phase cyclin-dependent protein kinases (CDK), this pre-replicative-complex (pre-RC, consisting of ORC, Cdc6, Cdt1, MCM proteins and possibly several additional factors) is likely to trigger initiation of DNA replication (reviewed in Dutta and Bell, 1997, Lei and Tye, 2001). Recent reports showed that *ori* in budding and fission yeast have a single precise start site adjacent to the ORC binding site from where DNA replication begins (Bielinsky and Gerbi, 1999, Gomez and Antequera, 1999).

After origin DNA has become licensed for replication, a first step would seem to be the local unwinding of DNA around the origin, resulting in a 'replication bubble' (see Figure 1.1, bottom panel). In *in vitro* replication of SV40 DNA, T-antigen incorporates two functions: it serves as an initiator protein and acts as a helicase, opening up the DNA (reviewed in Hassell and Brinton, 1996). The identity of the

unwinding helicase in eukaryotic cells is not yet clear although evidence accumulates that the six-associating-MCM proteins could carry out the helicase function (Ishimi, 1997, Labib *et al.*, 2000, reviewed in Labib and Diffley, 2001). In the SV40 system, the single-stranded-binding protein RPA coats the helicase-unwound DNA and stabilises it. Furthermore, it stimulates the helicase activity of SV40 T-antigen and interacts with the DNA polymerase α (pol α)/primase-complex which subsequently primes DNA synthesis (see below and reviewed in Waga and Stillman, 1998). Figure 1.1 gives a representation of the initiation events of DNA synthesis.

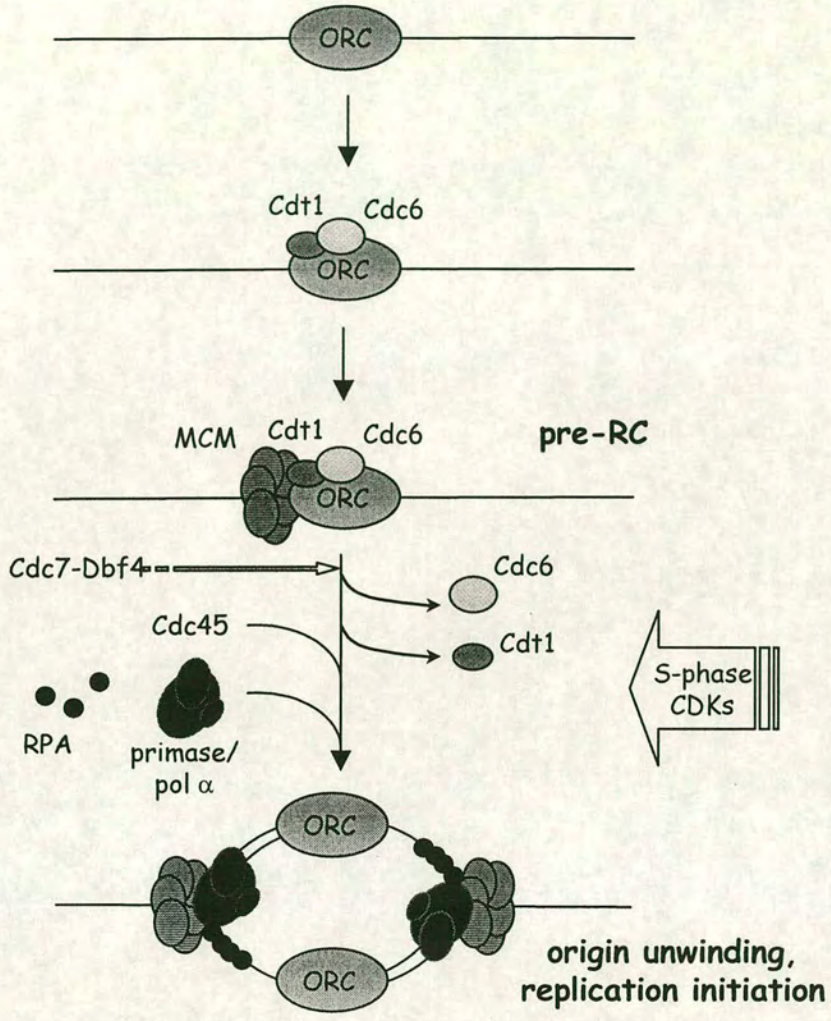



**Figure 1.1**: Schematic representation of initiation of yeast DNA replication. ORC is bound to origin DNA throughout the cell cycle. In G1 assembly of the pre-RC begins with binding of Cdc6 and Cdt1 to ORC and is complete with subsequent recruitment of the MCM proteins. Following phosphorylation of the MCM-complex during S-phase by Cdc7-Dbf4 kinase and removal of Cdc6 and Cdt1 other replication factors, like Cdc45, RPA and pol α/primase, triggers initiation of DNA synthesis. See text for further details.

### 1.1.3 Elongation

After initiation of DNA replication at the *ori*, DNA synthesis proceeds bidirectionally with replication forks moving away from the origin in either direction. The leading strands are synthesised continuously in the 5'→3' direction whereas the lagging strands are generated discontinuously through the formation of Okazaki fragments. Interestingly, the two strands are synthesised co-ordinately (Kornberg and Baker, 1992, Lee *et al.*, 1998).

DNA synthesis of the leading strand as well as that of each Okazaki fragment on the lagging strand is catalysed by the primase subunits of pol α/primase, which generate a short RNA primer that is then elongated with deoxynucleotides by pol α (Waga *et al.*, 1994). Pol α/primase is the only protein complex capable of synthesising DNA *de novo*. Since pol α/primase is a nonprocessive enzyme and has no proofreading activity, bulk DNA synthesis is carried out by another DNA polymerase, namely pol δ (and/or pol ε, see below). The 3' termini of the RNA-DNA primers, which are also called initiator DNAs (iDNA), are recognised by RF-C (Tsurimoto and Stillman, 1990) which then promotes a process called polymerase switching (Tsurimoto and Stillman, 1991b). The *clamp-loader* RF-C binds at the primer-template junction and loads PCNA (the *sliding clamp*) onto the DNA in an ATP-dependent manner, thereby displacing pol α/primase (Tsurimoto and Stillman, 1991b). PCNA, encircling the DNA, in turn then tethers the replicative DNA polymerase δ (or pol ε) to the site of DNA synthesis (Tsurimoto and Stillman, 1991a), completing the assembly of the DNA polymerase δ holoenzyme (which consists of pol δ, RF-C and PCNA). Through association with PCNA, DNA polymerase δ (or pol ε) is able to synthesise the leading and lagging strand highly processively (Prelich *et al.*, 1987b; Bauer and Burgers, 1988a). The RF-C-mediated polymerase switch occurs during the priming of the leading strand and during synthesis of every Okazaki fragment (Tsurimoto and Stillman, 1991b; Waga and Stillman, 1994; reviewed in Waga and Stillman, 1998). Recent *in vitro* binding experiments suggest that RPA plays a crucial role in triggering the polymerase switch by serving as a binding partner at the replication fork which pol α, RF-C and pol δ compete for (Yuzhakov *et al.*, 1999). On the lagging strand DNA synthesis of the Okazaki fragment continues at least until

the polymerase encounters the preceding Okazaki fragment. The final processing of Okazaki fragments, which is the subject of the next section, results in the formation of a continuous DNA strand.
Figure 1.2 summarises the priming and elongation steps of DNA replication.

The role of pol ε in DNA replication is not yet clear. Although dispensable for *in vitro* replication of SV40 DNA by human cell extracts (Lee *et al.*, 1991b) it is required for DNA replication in budding and fission yeast (Budd and Campbell, 1993; d'Urso and Nurse, 1997). However, recent reports showed that in budding yeast the essential function of the pol ε homologue resides in the C-terminal domain which does not contain the polymerase and 3'→5' exonuclease motifs (Dua *et al.*, 1999, Kesti *et al.*, 1999). On the other hand, earlier studies have shown that the exonuclease function of pol ε is active during *S. cerevisiae* DNA replication (Morrison and Sugino, 1994, Shcherbakova and Pavlov, 1996) suggesting that maybe pol δ can substitute for non-functional pol ε. It has been suggested that pol δ and pol ε replicate different strands (Karthikeyan *et al.*, 2000) or that pol ε may be involved in early steps of DNA replication since it has been shown that pol ε dissociates from origin DNA and associates with non-origin DNA during S phase (Aparicio *et al.*, 1997). Pol δ, on the other hand, may function during Okazaki fragment processing since it interacts with proteins involved in this process (Kang *et al.*, 2000). More work will be necessary to establish the precise roles of DNA polymerases δ and ε during DNA replication.

Further understanding of details of the events that are taking place during eukaryotic DNA replication may in the future come from the study of archaebacteria. Their replication proteins are generally more closely related to eukaryotes than eubacteria, but fewer proteins seem to perform the same functions (reviewed in MacNeill, 2001a).

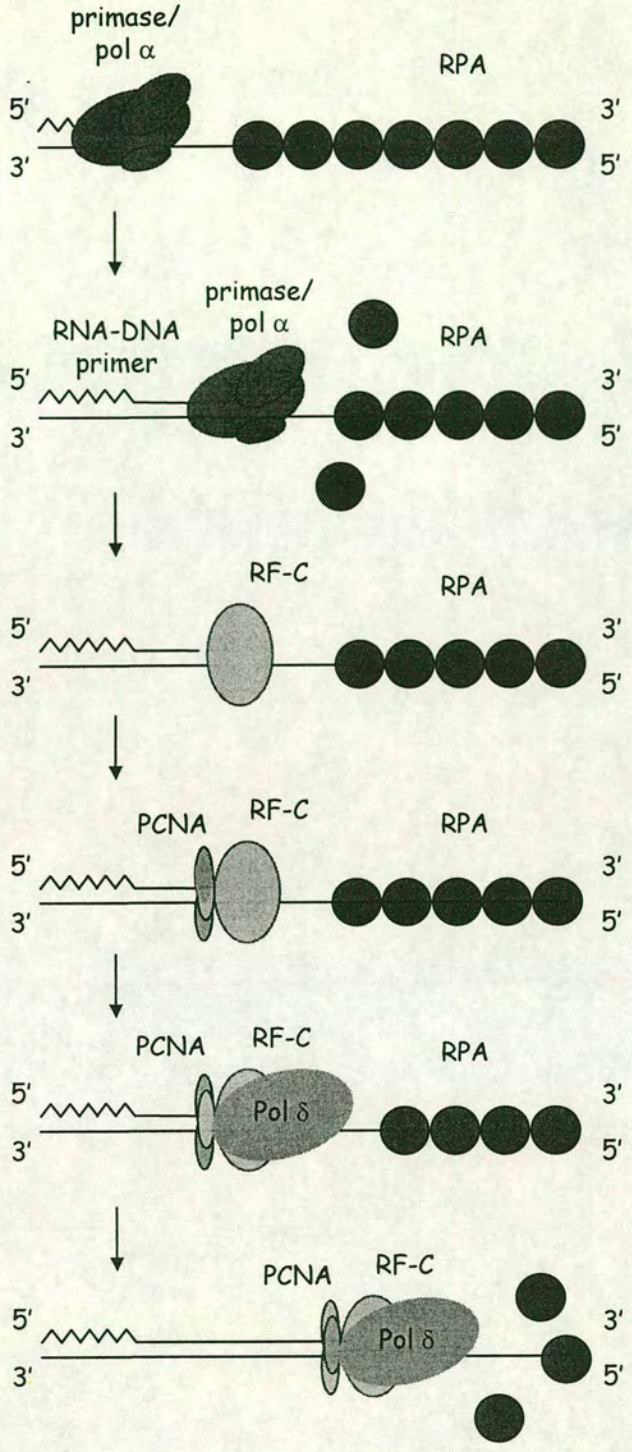



**Figure 1.2**: Schematic representation of polymerase switching. The RNA-DNA primer is synthesised by pol α/primase. RF-C binds to the 3' end of the initiator DNA, thereby replacing pol α/primase, and loads PCNA onto the DNA. This is followed by recruitment of DNA polymerase δ and subsequent processive DNA synthesis. Single stranded DNA is coated by RPA. See text for further details.

## 1.1.4 Okazaki fragment processing

### 1.1.4.1 Introduction

During DNA replication the lagging strand is synthesised discontinuously through a series of discrete Okazaki fragments, which is a consequence of the unique 5' → 3' directionality of the replicative polymerases. The synthesis of these short DNA stretches as intermediates in DNA replication were first discovered in prokaryotes (Okazaki *et al.*, 1968) and later in mammalian, viral and fungal systems (reviewed in Edenberg and Huberman, 1975, Nasmyth, 1977). In prokaryotes Okazaki fragments are ~1000 nucleotides in size, whereas in eukaryotes Okazaki fragments are only 50-150 nucleotides long (reviewed in Edenberg and Huberman, 1975).

Like the leading strand, each Okazaki fragment is initiated by primase synthesising a short stretch of around 10 ribonucleotides which is extended by pol α for a further ~30 deoxyribonucleotides (Waga and Stillman, 1998). Since pol α lacks proofreading activity, which ensues in primed DNA being error-prone, it needs to be removed in a process termed Okazaki fragment maturation or processing. The efficient maturation of Okazaki fragments is vital for maintaining a cell's genomic integrity, requiring a highly co-ordinated interplay between the participating proteins. For example yeast have a genome size of ~15 Mb and Okazaki fragments are <150 nucleotides long. Consequently, around 100,000 Okazaki fragments are produced during S phase – which lasts only ~20 minutes. Each of these must be processed. Okazaki fragment maturation involves RNA-DNA primer removal, error-free replacement synthesis and sealing of the resultant nick between adjacent Okazaki fragments.

The next section describes DNA replication proteins involved in Okazaki fragment processing in addition to PCNA and pol δ and the following section introduces a current model of the mechanisms entailed in the maturation of Okazaki fragments.

### 1.1.4.2 FEN1, RNase HI and Dna2

Through analysis of SV40 DNA replication, FEN1 and RNase HI were identified as having necessary activities during DNA replication (Waga *et al.*, 1994, Waga and Stillman, 1994) and it is likely that they are required for maturation of Okazaki fragments (see next section). In conjunction with a DNA polymerase and DNA ligase I these two factors were capable of completing lagging strand replication on model Okazaki fragment substrates (Turchi *et al.*, 1994). FEN1 is a structure-specific endonuclease which cleaves at branched DNA structures that have a single stranded 5' flap. In addition, it has exonucleolytic activity acting primarily at a nick (reviewed in Lieber, 1997). RNase HI is a ribonuclease which has RNA-DNA junction-specific RNase activity (Murante *et al.*, 1998). The fact that yeast strains deleted for the FEN1 and/or RNase HI homologues Rad27p and RNase H(35), respectively, are still viable – although the phenotype of a *rad27Δ* strain is consistent with a role for FEN1 in DNA replication – suggests the existence of another nucleolytic activity capable of assisting in the removal of RNA-DNA primers (Sommers *et al.*, 1995, Qiu *et al.*, 1999).

The Dna2 enzyme which comprises DNA helicase and single-stranded DNA endonuclease activities seems a suitable candidate. Dna2 was discovered as a protein in *S. cerevisiae* that is essential for DNA replication (Budd and Campbell, 1995) and it functions late during S phase (Fiorentino and Crabtree, 1997). Dna2 is essential for viability in the yeasts *S. pombe* and *S. cerevisiae* (Kang *et al.*, 2000, Budd and Campbell, 1995) and is conserved in humans, *A. thaliana*, *X. laevis* and *C. elegans* (Eki *et al.*, 1996, Liu *et al.*, 2000). Recent studies have identified the endonuclease activity of Dna2 to be essential for viability, whereas the helicase activity is not (Lee *et al.*, 2000, Budd *et al.*, 2000). Strong indications that Dna2 may function in maturation of Okazaki fragments came from the finding that the yeast protein interacts genetically and biochemically with the FEN1 homologue Rad27p (Budd and Campbell, 1997, Kang *et al.*, 2000). Furthermore, yeast Dna2 genetically interacts with the DNA polymerase δ B- and C-subunits, DNA ligase I and each of the three subunits of the single-stranded binding protein RPA (Kang *et al.*, 2000, Bae *et al.*, 2001).

### 1.1.4.3 A model for maturation of Okazaki fragments

Recent *in vitro* studies on Dna2p and other replication factors in *S. cerevisiae* gave novel insights into the mechanisms of Okazaki fragment maturation in yeast, from which a new model emerged (Bae and Seo, 2000, Bae *et al.*, 2001). In this model, when pol δ/PCNA encounter the adjacent Okazaki fragment, it carries out displacement synthesis of the downstream RNA-DNA primer thereby forming a 5' RNA-DNA flap structure. When the nascent flap structure reaches a critical size, RPA forms a complex with the DNA portion of the primer and recruits Dna2 onto the flap. At this point Dna2 may contact pol δ subunits which could possibly result in a cessation of displacement synthesis. Dna2 cleaves off the RNA portion of the initiator DNA. Dissociation of the cleaved iDNA, RPA, Dna2 and pol δ leaves behind PCNA, encircling the double-stranded DNA at the site of the shortened flap. PCNA can then tether FEN1 to the remaining flap structure and stimulate its endonucleolytic activity to produce a nick between two neighbouring Okazaki fragments. DNA ligase I then seals this nick to produce a continuous DNA strand.

This model justifies several observations. Firstly, the prerequisite that RPA must coat the primed DNA before Dna2 action can occur could account for continuation of displacement synthesis by pol δ until the erroneous RNA-DNA primer is completely removed. In that way accurate duplication of the primer region by pol δ is assured. The sequential action of Dna2 before FEN1 could explain why Dna2 activity is an essential step which creates the substrate for FEN1. In the absence of FEN1 other members of the FEN1-family of structure-specific nucleases may substitute for its function. This model does not include RNase HI, however. Either RNase HI functioning together with FEN1 is part of an alternative or parallel minor pathway of Okazaki fragment processing or it plays a more important role in Okazaki fragment processing of higher eukaryotes.

The above described model awaits further *in vivo* studies to support its validity but, so far, provides the most convincing interpretation of the gathered evidence. Figure 1.3 gives an outline of the model for processing Okazaki fragments as suggested by Seo and co-workers and reviewed in MacNeill, 2001b.

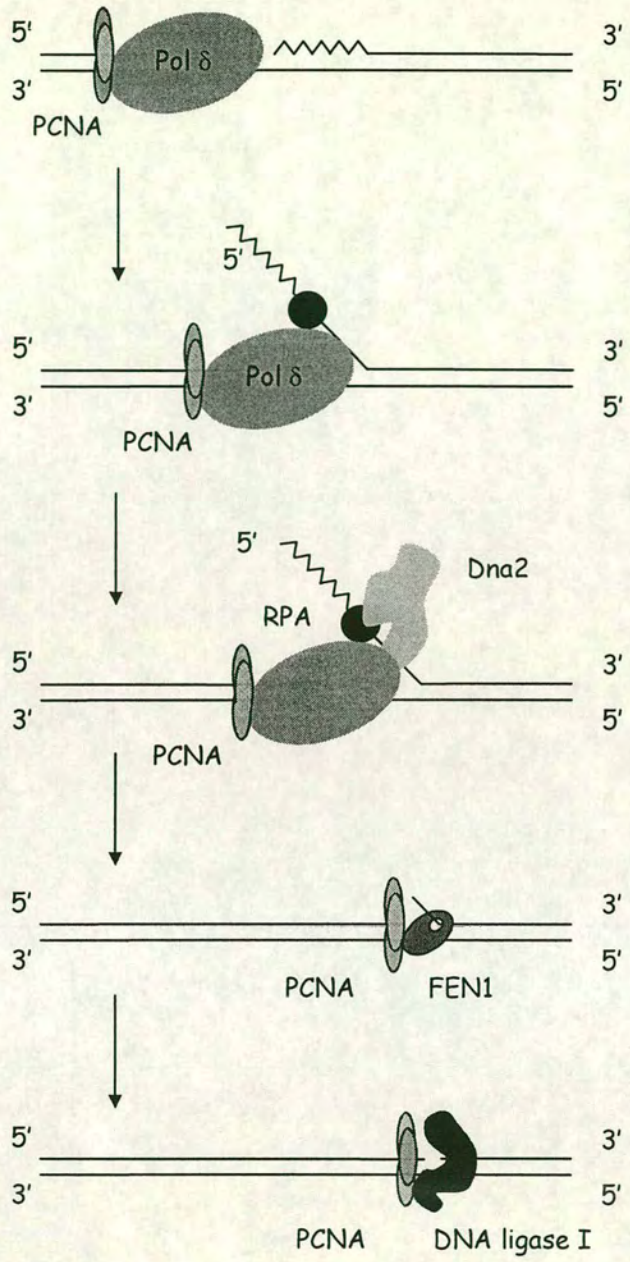



**Figure 1.3**: A model for Okazaki fragment processing according to the Seo lab (see text for references). The RNA portion of the previous Okazaki fragment is denoted as a zigzag line. Displacement synthesis by pol δ generates a flap structure and renders the primer single-stranded. RPA binds to the primer DNA, recruits Dna2 which then cleaves off the primer RNA. Dissociation of pol δ, Dna2 and RPA allows the association of FEN1 with PCNA. FEN1 cleaves the flap junction and finally DNA ligase I seals the resultant nick between adjacent Okazaki fragments.

## 1.1.5 PCNA as a key player in DNA replication

### 1.1.5.1 Introduction

PCNA plays various roles in DNA metabolism being involved in DNA replication, DNA repair, recombination, cell cycle regulation and checkpoint control.

It was originally identified by two different labs. Miyachi *et al.* (1978) discovered an auto-antigen in patients with lupus erythematosus. Since this protein was observed in nuclei of dividing cells it was named Proliferating Cell Nuclear Antigen (PCNA). Two years later it was discovered as a protein that is synthesised in S-phase (Bravo and Celis, 1980). PCNA is required for SV40 replication *in vitro* (Prelich *et al.*, 1987a; Prelich *et al.*, 1987b) and one major role that it serves is as a processivity factor for DNA polymerase δ (Tsurimoto and Stillman, 1989) during DNA replication (for review see Kelman, 1997).

Genes encoding PCNA have been cloned and sequenced from a variety of species including human (Almendral *et al.*, 1987), *S. cerevisiae* (Bauer and Burgers, 1990) and *S. pombe* (Waseem *et al.*, 1992). Although there is only moderate sequence similarity - 32% - between yeast and mammalian PCNA, they can substitute for one another in *in vitro* replication assays with pol δ (Bauer and Burgers, 1988a and 1988b) and human PCNA can complement an *S. pombe* PCNA deletion (Waseem *et al.*, 1992) confirming that they are functional homologues.

PCNA is an essential protein. Expression of PCNA anti-sense RNA in exponentially growing cells causes a suppression of DNA replication and cell cycle progression (Jaskulski *et al.*, 1988) and deletion of the PCNA gene in both *S. cerevisiae* and *S. pombe* is lethal (Bauer and Burgers, 1990; Waseem *et al.*, 1992).

The crystal structure of yeast and human PCNA has been determined (Krishna *et al.*, 1994; Gulbis *et al.*, 1996). PCNA is a ring-shaped molecule that can encircle double stranded DNA. It forms a stable homotrimer in solution. The trimeric ring has an overall six-fold symmetry as each monomer consists of two globular domains of related structure. The two domains, which are connected by an extensive loop, called the inter-domain connector loop, are structurally almost identical but show no sequence homology. The charge distribution in the molecule is highly asymmetric; the outer surface is very negatively charged whereas there is a net positive

electrostatic potential along the central channel. Eukaryotic PCNA is structurally homologous to the *E. coli* β sliding clamp and the T4 gene 45 protein. Although the β sliding clamp is a homodimer its three dimensional structure is almost exactly superimposable on PCNA as each monomer consists of three related domains (Kelman and O'Donnell, 1995; for review see Kelman, 1997).

In the next two sections the *clamp-loader* RF-C and the replicative DNA polymerase δ enzyme are described briefly and what is known about their interactions with the PCNA *sliding clamp*. The third section deals with proteins involved in DNA metabolism and cell cycle progression which bind to PCNA via a conserved PCNA binding motif. A short review of what is known about the interaction between PCNA and several DNA replication proteins follows.

### 1.1.5.2 The *clamp-loader* replication factor C

RF-C, also called activator 1, was first identified as an essential factor for *in vitro* SV40 DNA replication (Tsurimoto and Stillman, 1989; Lee *et al.*, 1991a). Its main role is to load the DNA polymerase δ processivity factor PCNA onto the DNA.

RF-C is a five subunit complex consisting of one large subunit -140 kDa in the human and ~95 kDa in yeast, and four small subunits - 36 kDa to 40 kDa in both organisms. The subunits of those two species share between 36 % (large subunit) and 60 % (hRF-C40 and ScRF-C4, respectively) sequence homology. Genes of all five subunits of *S. cerevisiae* and *S. pombe* have been identified (reviewed in MacNeill and Burgers, 2000). In yeast, all five subunits are essential (Cullmann *et al.*, 1995) and are all required to form a stable complex (Uhlmann *et al.*, 1996). Studies in both yeasts identified a role for some RF-C subunits in DNA replication checkpoint function (Sugimoto *et al.*, 1996, Reynolds *et al.*, 1999). Recently, the crystal structure of the *E. coli* clamp-loader γ complex has been solved which offers insights into the precise mechanism of clamp loading (Jeruzalmi *et al.*, 2001) and the crystal structure of a small subunit of the archaeal *clamp-loader* from *Pyrococcus furiosus*

indicates a similar structure (Oyama *et al.*, 2001, reviewed in O'Donnell *et al.*, 2001).

The five different subunits have seven different regions of protein sequence similarity in common, termed RF-C boxes II to VIII. The large subunit contains an additional box, RF-C box I, that shares a region of high similarity both with prokaryotic DNA ligases and to a lesser extent with poly(ADP-ribose)polymerases, which contain a version of the BRCT motif. A superfamily of proteins sharing this motif seems to be involved in DNA-damage responsive checkpoints. The DNA-binding region maps within this RF-C box I (Fotedar *et al.*, 1996). A PCNA-binding domain that is conserved in all five subunits of RF-C has been identified downstream of the DNA-binding domain (Fotedar *et al.*, 1996). Deletions into RF-C box II, which maps within the PCNA binding domain, in any of the subunits have been shown to disrupt the DNA replication function, i.e. loading PCNA onto DNA (Uhlmann *et al.*, 1997).

### 1.1.5.3 DNA Polymerase δ

The well-established role for PCNA is to serve as a processivity factor for DNA polymerase δ (Prelich *et al.*, 1987b). Pol δ possesses 5'→3' DNA polymerase activity and 3'→5' exonuclease activity (Wang, 1991). In addition to DNA polymerase α, pol δ is essential for DNA replication *in vitro* in the SV40 system and in budding yeast (Melendy and Stillman, 1991; Budd and Campbell, 1993). In mammalia and yeast, pol δ comprises at least three subunits – a large catalytic subunit (A-subunit) and two or more smaller subunits (B-, C- and D-subunit) (reviewed in MacNeill *et al.*, 2001).

The question of which subunit of pol δ interacts with PCNA resulting in the stimulation of polymerase activity is not solved unambiguously as yet. The catalytic subunits of budding yeast and human pol δ on their own are not stimulated by PCNA (Tratner *et al.*, 1997) although there have been studies reporting the direct interaction of the mammalian catalytic subunit with PCNA (Zhang *et al.*, 1995, Zhang *et al.*,

1999). A heterodimer consisting of the catalytic subunit and the B-subunit is found capable of interacting with PCNA resulting in processive DNA synthesis *in vitro* (Burgers and Gerik, 1998; Zhou *et al.*, 1997). In contrast, other recent studies show that the formation of an active pol δ complex, which can be stimulated by PCNA requires the presence of the C-subunit as well (Ducoux *et al.*, 2001, Shikata *et al.*, 2001, see also next section). Several lines of evidence support the existence of several distinct binding sites for PCNA within the pol δ complex. It remains to be seen what precise mechanisms underlie the stimulation of pol δ activity and if there are additional functions to any pol δ – PCNA interactions.

### 1.1.5.4 PCNA binding via a conserved motif

In the past few years PCNA has been found to interact with numerous proteins that function in DNA replication, DNA repair and cell cycle regulation. This implies a central role for PCNA in linking the diverse protein functions acting at the replication fork and in serving as a co-ordinator between DNA metabolic events and cell cycle progression. The PCNA-interacting proteins that participate in DNA replication and in at least one pathway of DNA repair, are listed in Table 1.1.

**Table 1.1**: PCNA-interacting-proteins that are involved in DNA replication, DNA repair and cell cycle progression. Most references are listed in Tsurimoto, 1999 and Warbrick, 2000. [1] Niimi *et al.*, 2001, [2] Kleczkowska *et al.*, 2001, [3] Tsuchimoto *et al.*, 2001

| DNA replication & modification | DNA repair | cell cycle progression |
|---|---|---|
| DNA polymerase δ | XP-G | p21 |
| RF-C | MSH2 | CDK/cyclin-complexes |
| FEN 1 | MLH1 | |
| DNA ligase I | MSH3[2] | |
| *S. pombe* Cdc24 | MSH6[2] | |
| 5-methyltransferase | uracil DNA glycosylase | |
| DNA topoisomerase II[1] | AP endonuclease[3] | |
| | WRN helicase? | |

Many studies have been carried out to identify the regions of PCNA involved in these protein interactions, and naturally the areas within these
proteins that interact with PCNA in return (reviewed in Warbrick, 2000). In a recent report, polymerase δ, RF-C, FEN1, p21 and DNA ligase I were all found to bind to part of the inter-domain connector loop and loops on one face of the trimer close to the C-terminus (Jonsson *et al.*, 1998). Strikingly, the proteins share versions of a conserved amino acid sequence motif – Qxx(I/L)xxFF - which has been shown to be at least partly responsible for the interaction with PCNA (Warbrick, 1998). This motif was first identified in the protein kinase inhibitor p21 and is hence named the '$p21^{CIP1}$-like motif'. Figure 1.4 shows the structure of PCNA bound to a p21 peptide and in Figure 1.5 the sequences around the conserved PCNA-binding motifs of several proteins from different species are aligned. Mutational analyses of several proteins have shown that alteration of the isoleucine and/or phenylalanine residues abolishes PCNA-binding *in vitro* (Gary *et al.*, 1997, Montecucco *et al.*, 1998). As part of this work, the presence of the PCNA binding motif in the fission yeast DNA ligase I protein for *in vivo* function will be assessed (Chapter 4)
The role of PCNA in DNA replication and its interaction with various proteins is reviewed in Tsurimoto, 1999.

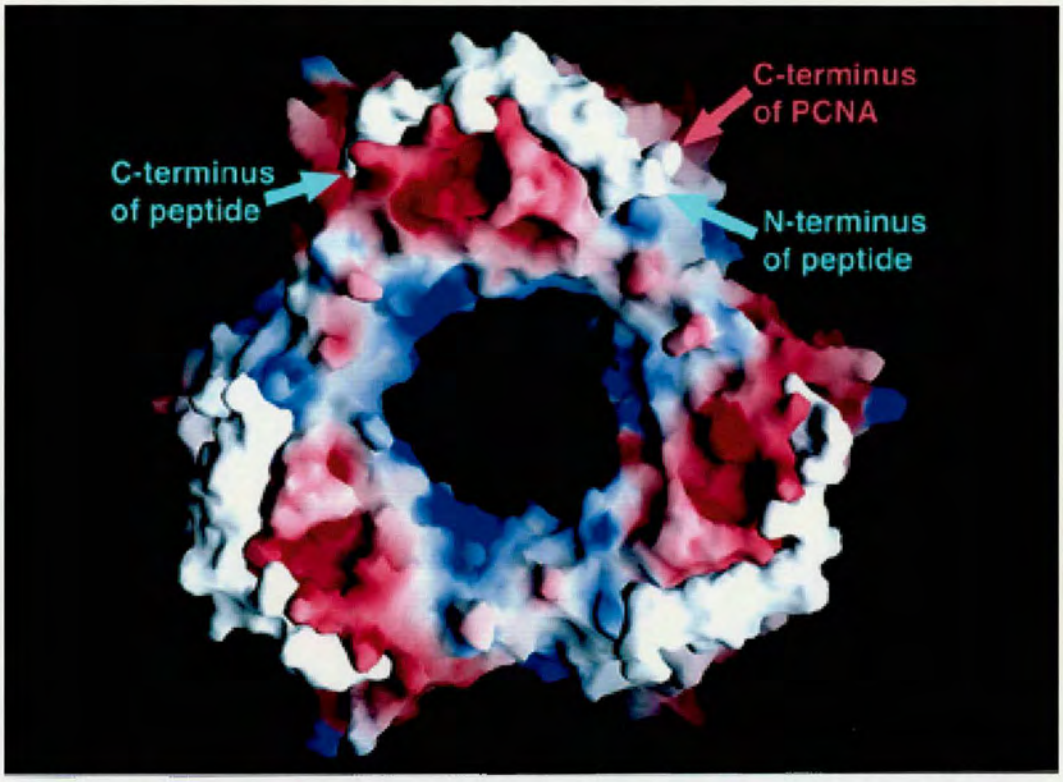



**Figure 1.4**: Structure of human PCNA bound to a p21 peptide (white). PCNA is coloured according to its electrostatic potential, with regions of intense positive charge appearing blue and electronegative regions shown in red. Image taken from Gulbis *et al.*, 1996.

**Figure 1.5**: Sequence alignment of $p21^{CIP1}$-like PCNA binding motifs of a selection of proteins from different species. Conserved residues of the motif Qxx(I/L)xxFF are shown in bold and its positions within the sequence are boxed. The numbers denote the amino acid position of the first shown residue in the protein and an asterisk indicates the C-terminal end of the protein.

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Ligase I** | | | | | | | | | | | | | | | | | | | | | |
| human | 1 | | | M | **Q** | R | S | **I** | M | S | **F** | **F** | H | P | K | K | E | G | K | A | K |
| *Xenopus* | 1 | | | M | **Q** | R | T | **I** | K | S | **F** | **F** | Q | P | K | L | G | A | E | V | K |
| *S. pombe* | 21 | S | T | R | **Q** | S | D | **I** | S | N | **F** | **F** | I | S | S | A | S | H | K | S | E |
| *S. cerevisiae* | 35 | K | P | K | **Q** | A | T | **L** | A | R | **F** | **F** | T | S | M | K | N | K | P | T | E |
| **Fen1** | | | | | | | | | | | | | | | | | | | | | |
| human | 334 | G | S | T | **Q** | G | R | **L** | D | D | **F** | **F** | K | V | T | G | S | L | S | S | A |
| *Xenopus* | 334 | G | S | T | **Q** | G | R | **L** | D | D | **F** | **F** | K | V | T | G | S | I | S | S | T |
| *S. pombe* | 335 | T | I | P | **Q** | G | R | **L** | D | S | **F** | **F** | K | P | V | P | S | S | P | K | K |
| *S. cerevisiae* | 334 | S | G | I | **Q** | G | R | **L** | D | G | **F** | **F** | Q | V | V | P | K | T | K | E | Q |
| **XPG** | | | | | | | | | | | | | | | | | | | | | |
| human | 987 | Q | Q | T | **Q** | L | R | **I** | D | S | **F** | **F** | R | L | A | Q | Q | E | K | E | D |
| *Xenopus* | 1020 | Q | Q | T | **Q** | L | R | **I** | D | S | **F** | **F** | R | L | E | Q | H | E | A | A | G |
| *S. pombe* | 980 | V | G | T | **Q** | S | N | **L** | T | Q | **F** | **F** | E | G | G | N | T | N | V | Y | A |
| *S. cerevisiae* | 992 | K | G | K | **Q** | K | R | **I** | N | E | **F** | **F** | P | R | E | Y | I | S | G | D | K |
| **p21** | | | | | | | | | | | | | | | | | | | | | |
| human | 141 | K | R | R | **Q** | T | S | M | T | D | **F** | Y | H | S | K | R | R | L | I | F | S |
| mouse | 136 | K | R | R | **Q** | T | S | **L** | T | D | **F** | Y | H | S | K | R | R | L | V | F | C |
| *Drosophila* | 181 | R | K | R | **Q** | P | K | **I** | T | E | **F** | M | K | E | R | K | R | L | A | Q | A |
| **MCMT** | | | | | | | | | | | | | | | | | | | | | |
| human | 40 | S | T | R | **Q** | T | T | **I** | T | S | H | **F** | A | K | G | P | A | K | R | K | P |
| mouse | 158 | T | T | R | **Q** | T | T | **I** | T | H | A | **F** | T | K | G | P | T | K | R | K | P |
| *Xenopus* | 149 | A | G | K | **Q** | P | T | **I** | L | S | M | **F** | S | K | G | S | T | K | R | K | S |
| **Pol δ C-subunit** | | | | | | | | | | | | | | | | | | | | | |
| human | 476 | A | N | R | **Q** | V | S | **I** | T | G | **F** | **F** | Q | R | K | * | | | | | |
| *S. cerevisiae* | 336 | L | K | K | **Q** | G | T | **L** | E | S | **F** | **F** | K | R | K | A | K | * | | | |
| *S. pombe* | 357 | K | P | Q | **Q** | K | S | **I** | M | S | **F** | **F** | G | K | K | * | | | | | |
| **RFC (large subunit)** | | | | | | | | | | | | | | | | | | | | | |
| human | 1 | | | | | M | D | **I** | R | K | **F** | **F** | G | V | I | P | S | G | K | K | L |
| *Drosophila* | 1 | | | M | **Q** | R | G | **I** | D | S | **F** | **F** | K | R | L | P | A | K | A | K | S |
| *S. cerevisiae* | 1 | | | | M | V | N | **I** | S | D | **F** | **F** | G | K | N | K | K | S | V | R | S |
| *S. pombe* | 1 | | M | S | N | S | D | **I** | R | S | **F** | **F** | G | G | G | N | A | Q | K | K | P |

RF-C contains distinct sites that mediate an interaction with PCNA. The central region of the large subunit Rfc1, which is homologous to the N-terminal portions of Rfc2-Rfc5 and contains RF-C boxes II – VIII, is involved in an interaction with PCNA and its loading onto the DNA (Fotedar *et al.*, 1996, see above).

In collaboration between different labs, another region within the large subunit of human RF-C that interacts with PCNA was identified last year (Montecucco *et al.*, 1998). The extreme N-terminus of hRF-Cp140 contains a version of the p21-like PCNA binding motif and an N-terminal fragment binds PCNA (Montecucco *et al.*, 1998). The precise function of this interaction is not yet clear since a large N-terminal deletion of human or yeast Rfc1 including the PCNA binding motif did not negatively affect the replication function of the RF-C complex in *in vitro* assays (Uhlmann *et al.*, 1997, Gomes *et al.*, 2000). One possible explanation for interaction with PCNA via this motif came from the observation that the motif of human Rfc1, together with a stretch of basic amino acids, which could function as a nuclear localisation signal, localised to sites of ongoing DNA replication (Montecucco *et al.*, 1998).

PCNA constitutes the processivity factor for DNA polymerase δ converting it from a distributive to a processive mode of action. Although contradictory reports exist (see previous section) recent evidence suggests that the crucial interaction between the pol δ complex and PCNA may well be mediated via the C-subunit. Pol32p, Cdc27 and p66 show limited sequence similarity but they share the conserved $p21^{CIP1}$-like PCNA-binding motif in their C-termini. All three proteins bind PCNA (Gerik *et al.*, 1998; Reynolds *et al.*, 2000, Hughes *et al.*, 1999). The assignment of an important function for the C-subunit, however, is tempered by the fact that Pol32p is not essential for cell viability in *S. cerevisiae* (Gerik *et al.*, 1998). On the other hand, fission yeast C-subunit Cdc27 is essential (MacNeill *et al.*, 1996) and the PCNA binding motif of Cdc27, which is the C-subunit of *S.pombe* pol δ, is required for protein function *in vivo* (Reynolds *et al.*, 2000). In keeping with the latter finding are recent studies which found the presence of the C-subunit necessary for PCNA-dependent stimulation of human DNA polymerase δ activity (Ducoux *et al.*, 2001, Shikata *et al.*, 2001). The contradicting results mentioned here could be explained by

the presence of several distinct binding sites which had differential affinities for PCNA within the complex which contributed to stimulation of pol δ activity.

The most detailed studies carried out to date on the interaction with PCNA via the conserved motif have been carried out with FEN1. PCNA has been shown to interact with FEN1 and stimulate its endo- and exonucleolytic activities up to 50-fold (Li *et al.*, 1995). The PCNA binding site in human FEN1 was shown to be in the C-terminus where the conserved PCNA binding motif is located (Gary *et al.*, 1997, Warbrick *et al.*, 1997) and deletion of the region or mutation of the critical phenylalanine residues abolished binding *in vitro*. The contribution of a functional PCNA binding motif to the PCNA-dependent stimulation of FEN1 endonucleolytic activity has been found to vary amongst the groups who carried out the studies. For example, single and multiple mutations in the PCNA binding motif of human FEN1 did not have any apparent effect in the ability of FEN1 to cleave a flap-structure annealed to a linearised oligomer (Frank *et al.*, 2001). On the other hand, PCNA was unable to stimulate the catalytic activities of human FEN1 deleted for the PCNA binding motif on a circular flap template (Stucki et al., 2001). Similar results have been obtained with studies of the yeast PCNA and FEN1 homologues Pol30p and Rad27p. In addition to an important role for the motif in Rad27p to Pol30p binding and stimulation of its nuclease activity, the authors demonstrated that the protein contacts differ slightly depending on whether they are present in solution or complexed on the DNA (Gomes and Burgers, 2000).

The purpose of multiple proteins interacting with PCNA through a conserved motif remains unclear. If all these interactions serve physiologically important functions, this raises the questions of whether different proteins can bind to PCNA at the same time – keeping in mind that the homotrimeric PCNA has potentially three protein-binding sites, and how the various interactions which occur during DNA replication are co-ordinated temporally and spatially.

## 1.2 Eukaryotic DNA ligases

### 1.2.1 Introduction

DNA ligases are critical enzymes in DNA metabolism. They catalyse the joining of breaks in duplex DNA. The first DNA joining activities were discovered in *E. coli* in the late sixties and subsequently in mammalia and plants (Newman and Hanawalt, 1968, Lindahl and Edelman, 1968, Okazaki *et al.*, 1968, reviewed in Lehman, 1974). A very important cellular event that requires the action of a DNA ligase is the joining of Okazaki fragments (lagging strand intermediates) as well as the ligation of adjacent replicons, which are generated during semi-continuous DNA replication (see section 1.1). Other processes dependent on DNA ligases include various DNA repair pathways and genetic recombination. Thus, DNA ligases are crucial for the maintenance of genomic integrity and stability.

This chapter describes the characteristics of the two existing families of DNA ligases and gives details about the variety of ATP-dependent enzymes. After brief summaries of what is known about human DNA ligase III and DNA ligase IV proteins, DNA ligase I and its homologues are characterised in greater depth.

### 1.2.2 Classification, catalytic mechanism and structural aspects

DNA ligases fall into two groups based on their co-factor specificities. Generally, eubacteria encode DNA ligases that require $NAD^+$ for activity whereas eukaryotic, viral and archaebacterial enzymes utilise ATP as a cofactor. Despite very limited sequence similarity between the ATP- and $NAD^+$-dependent DNA ligases (see below) their reaction mechanism is very similar. The catalytic mechanism is outlined in Figure 1.6. Briefly, the formation of a phosphodiester bond between 3'-hydroxyl- and 5'-phosphate-termini at nicked DNA can be divided into three distinct catalytic steps. In a first reaction the DNA ligase becomes activated through the adenylation of the catalytic site lysine residue resulting in the formation of a covalent enzyme-AMP complex. The adenyl-group donor is either ATP or $NAD^+$ and the leaving group pyrophosphate or nicotinamide mononucleotide, respectively. Secondly, the

AMP moiety is transferred from the ligase molecule to the 5' phosphate group at the nick. In a final step, the phosphodiester bond is formed with the concomitant release of AMP (reviewed in Lehman, 1974, Timson *et al.*, 2000).

In ATP-dependent ligases six co-linear sequence motifs (I-VI) have been identified five of which participate in making conserved contacts with the ATP molecule. These motifs are common to the nucleotidyl transferase superfamily, which also include RNA ligases, tRNA ligases and eukaryotic mRNA capping enzymes (reviewed in Shuman and Schwer, 1995). Many mutational analyses on the motifs of different ligase proteins have been carried out which contributed to clarification of the catalytic mechanism (for example Luo and Barany, 1996, Sriskanda and Shuman, 1998 a and b, reviewed in Timson *et al.*, 2000, Doherty and Suh, 2000)
Previously, sequence alignments between $NAD^+$- and ATP-dependent ligases failed to detect any significant sequence similarity apart from motif I which includes the active site lysine residue. Consequently it had been speculated that the two classes of DNA ligases have evolved independently (Shuman and Schwer, 1995). However, crystal structures of the ATP-dependent ligases of bacteriophage T7 (Subramanya *et al.*, 1996) and the eukaryotic *chlorella* virus PBCV-1 (Odell *et al.*, 2000), the related GTP-dependent mRNA capping enzyme of PBCV-1 (Håkansson *et al.*, 1997) and of the $NAD^+$-dependent ligases of *Bacillus stearothermophilus* (an N-terminal fragment) (Singleton *et al.*, 1999) and *Thermus filiformis* (Lee *et al.*, 2000b) have now been solved and comparison of the structures of these three classes of enzymes revealed the conservation of a common 'core' structure, which consists of two domains with a deep cleft between them. Motifs I – V, situated in the N-terminal domain, line the groove between the domains and the active site is located at the base of the pocket (reviewed in Timson *et al.*, 2000, Doherty and Suh, 2000). The smaller C-terminal domain contains an OB-fold (oligonucleotide-binding), as determined from the ligase structures. Backed up by the available structural information the use of the recently developed iterative sequence database search PSI-BLAST (Altschul and Koonin, 1998) enabled Aravind and Koonin, 1999, to demonstrate a significant relationship between ATP- and $NAD^+$-dependent DNA ligases. Figure 1.7 shows a sequence alignment of the conserved motifs I-VI from selected members of the ATP-

and $NAD^+$-dependent DNA ligases as well as the GTP-dependent mRNA capping enzymes. Now it appears as if the two classes of DNA ligases may have originated from a common ancestor but that, while the core *structure* remained similar, the amino acid sequences diverged greatly – possibly to accommodate the accessibility to differing co-factors (Aravind and Koonin, 1999).

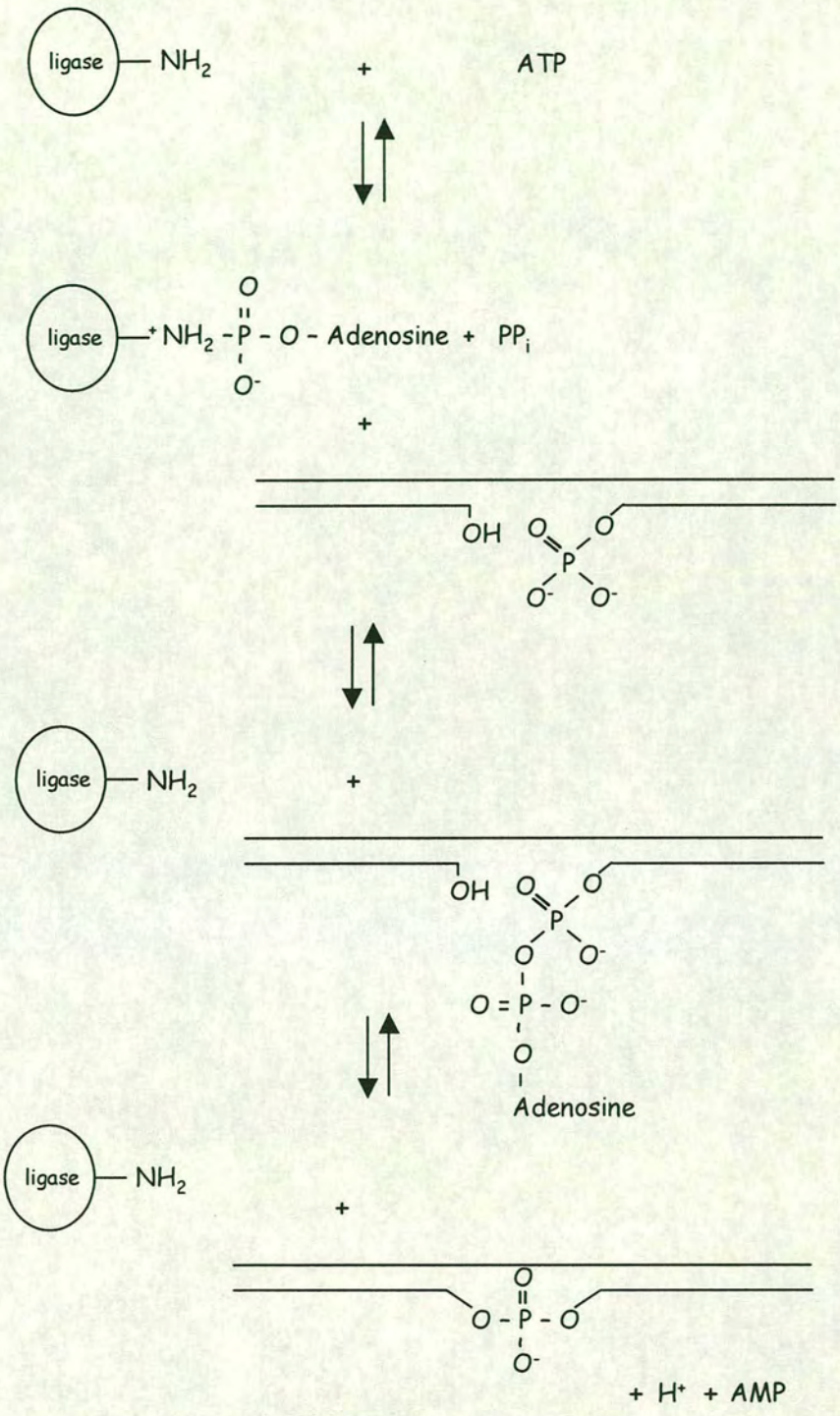



**Figure 1.6**: Catalytic mechanism of phosphodiester bond formation by ATP-dependent DNA ligases. See text for details.

| | | I | III | IIIa | IV | V | VI |
|---|---|---|---|---|---|---|---|
| | **Hs1 Lig** | KYDGQR | FILDTEAV | CLYAFDLIYLNG | EGLMV | WLKLKK--DYLDG | PRFIR--VREDK |
| | **Sp Lig** | KYDGER | FILDCEAV | CLFAFDILYLNG | EGLMV | WLKVKK--DYLSG | PRFIR--IREDK |
| | **Sc Lig** | KYDGER | LILDCEAV | CLFAFDILCYND | EGLMV | WLKLKK--DYLEG | PRFLR--IREDK |
| 1 | **Da Lig** | KYDGER | FIVEGEII | NVFLFDLMYYEG | EGVMV | WIKVKR--DYQSE | PRFIR--WRPDK |
| | **Vac Lig** | KYDGER | IVLDSEIV | CLFVFDCLYFDG | EGLVL | WLKIKR--DYLNE | PRFTR--IREDK |
| | **PBCV Lig** | KIDGIR | EGSDGEIS | SYYWFDYVT--- | EGVMI | LLKMKQ--FKDAE | PRFPV--FIGIR |
| | **T7 Lig** | KYDGVR | FMLDGELM | HIKLYAILPLHI | EGLIV | WWKMKP--ENEAD | PSFVM--FRGTE |
| | **Tfi Lig** | KVDGLS | LEVRGEVY | TFYALGLGLGLE | DGVVL | ALAYKF--PAEEK | PEVLR-VLKERR |
| 2 | **Bst Lig** | KIDGLA | LEARGEAF | DLFVYGLADAEA | DGIVI | AIAYKF--PAEEV | PEVVG-VVVDRR |
| | **Eco Lig** | KLDGLA | LEVRGEVF | TFFCYGVGVLEG | DGVVI | AVAFKF--PAQEQ | PQVVNVVLSERP |
| | **Sp CE** | KSDGIR | TLIDGELV | RYLVFDCLACDG | DGLIF | LLKWKPKEMNTID | WRFLR--FRDDK |
| 3 | **Sc CE** | KTDGLR | TLLDGELV | RYLMFDCLAING | DGLIF | LLKWKPEQENTVD | WEMLR--FRDDK |
| | **PBCV CE** | KTDGIR | SIFDGELC | AFVLFDAVVVSG | DGLII | LFKLKPGTHHTID | WKYIQ--GRSDK |

**Figure 1.7**: Sequence alignment of the conserved motifs in DNA ligases (Lig) and mRNA capping enzymes (CE). 1: ATP-dependent DNA ligases, 2: $NAD^+$-dependent ligases, 3: mRNA capping enzymes. Species names: HS1: human DNA ligase I, Sp: *Schizosaccharomyces pombe*, Sc: *Saccharomyces cerevisiae*, Da: *Desulfurolobus ambivalens*, Vac: *vaccinia* virus, PBCV: *Chlorella* virus, T7: bacteriophage T7, Tfi: *Thermus filiformis*, Bst: *Bacillus stearothermophilus*, Eco: *Escherichia coli*.

## 1.2.3 ATP-dependent ligases

### 1.2.3.1 Introduction

As mentioned above, ATP-dependent ligases are encoded by prokaryotic and eukaryotic viruses, archaea and eukaryotic species.

(There are, however, exceptions to this rule. For example, an ATP-dependent ligase was identified in the eubacterial obligate parasite *Haemophilus influenzae* (Cheng and Shuman, 1997) and recently an $NAD^+$-dependent ligase has been reported from a eukaryotic entomopoxvirus (Sriskanda *et al.*, 2001). A likely explanation would be that the presence of these untypical DNA ligases in those organisms was a result of horizontal gene transfer.)

All ATP-dependent ligases consist of a core, described in the previous section, which constitutes the minimal architecture necessary for phosphodiester bond formation. Enzymes from bacteriophages of the T series, i.e. T3, T4, T6 and T7 DNA ligase as well as the one from *Paramecium bursarium Chlorella* virus (PBCV) consist only of this core structure. The T7 enzyme can be seen as the archetype of a DNA ligase: it consists of 359 amino acids and has a molecular weight of only 41 kDa (see also Figure 1.8 b)). N-terminal to the core structure poxviral, archaeal and eukaryotic DNA ligase proteins have an additional domain, the precise function of which is not known (see below). Exclusive to eukaryotic DNA ligase proteins are further N- and/or C-terminal domains, which are likely to play regulatory roles or accommodate specialised functions. In accordance with this, in many eukaryotic species several ligase proteins have been identified. In mammalian cells three different ligase genes, *LIG1*, *LIG3* and *LIG4*, have been cloned so far and specific cellular functions have been assigned to their gene products (see following sections). They share the catalytic core and differ in their extensions. In Figure 1.8 a) the domain structure of human DNA ligase I, III and IV proteins are shown. Two additional ligase activities have been described in humans, designated ligase II and V. It is now clear that DNA ligase II is a proteolytic product of DNA ligase III and it remains to be seen if a new gene encoding for the DNA ligase V activity will be found (Husain *et al.*, 1995, Johnson and Fairman, 1997).

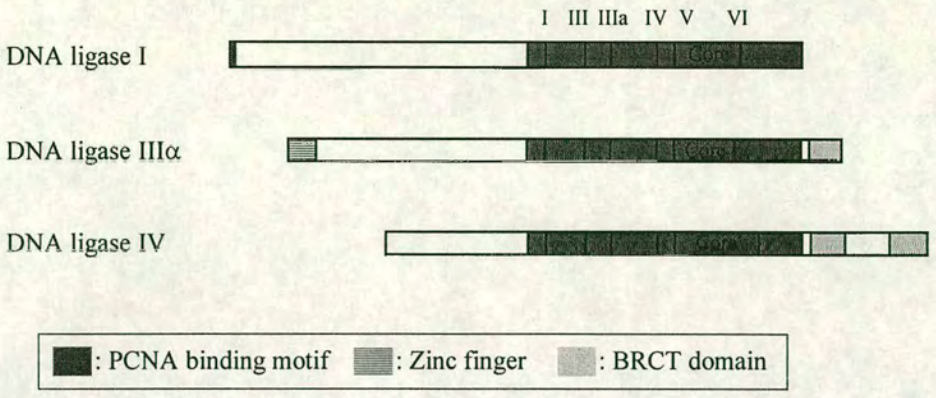



**Figure 1.8 a)**: Schematic representation of the domain structure of human DNA ligase proteins (for reviews see Tomkinson and Levin, 1997 and Tomkinson and Mackey, 1998)

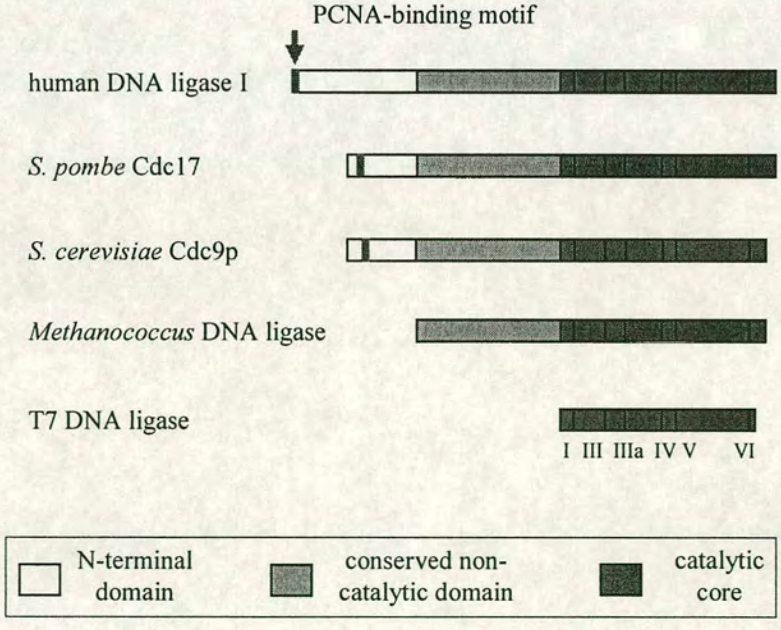



**Figure 1.8 b)**: Schematic representation of the domain structure of DNA ligase I proteins, the archaeal *Methanococcus* DNA ligase and the bacteriophage T7 DNA ligase.

**Table 1.2**: Properties of mammalian ligase activities (adapted from Tomkinson and Mackey, 1998). References are: [1] Barnes *et al.*, 1990, [2] Wei *et al.*, 1995.

| Mammalian ligase activity | I[1] | III[2] | IV[2] |
|---|---|---|---|
| gene | LIG1 | LIG3 | LIG4 |
| molecular mass (kDa) | 102 | 100 | 100 |
| substrates: | | | |
| oligo(dT)/ poly(dA) | + | + | + |
| oligo(dT)/ poly(rA) | - | + | + |
| oligo(rA) /poly(dT) | + | + | - |
| comments | | 2 splice variants: IIIα and IIIβ | |
| interacting proteins | PCNA, pol β | IIIα: XRCC1 | XRCC4 |
| Cellular process | replication, BER, mismatch repair?, NER? | IIIα: ss break repair, BER<br>IIIβ: DNA repair?, meiotic recombination? | NHEJ, V(D)J recombination |

Sequence comparisons and phylogenetic analyses indicate that archaeal enzymes as well as the yeast ligases Cdc9p (*S. cerevisiae*) and Cdc17 (*S. pombe*) are most closely related to mammalian DNA ligase I (Nakatani *et al.*, 2000). These enzmes display a relatively high degree of sequence similarity of a non-catalytic domain, located N-terminal to the catalytic domains. This domain will be designated conserved non-catalytic domain, or CNCD, in this work. Figure 1.8 b) gives an overview of the domain structure of DNA ligase I homologues in comparison to the archetype T7 DNA ligase. The characteristics of DNA ligase I proteins will be described in greater detail in section 1.2.4.

### 1.2.3.2 DNA ligase II and III

By the mid-seventies it became apparent that at least two distinct ligase acitivities were present in mammalia, DNA ligase I (see section 1.2.4) and DNA ligase II (reviewed in Soderhall and Lindahl, 1976). DNA ligase II, a proteolytic product of DNA ligase III, was subsequently purified from bovine thymus (Teraoka *et al.*, 1986). Not until another ten years later was DNA ligase III purified to near homogeneity from bovine testes (Husain *et al.*, 1995) and human DNA ligase III cDNAs were cloned and sequenced (Chen *et al.*, 1995, Wei *et al.*, 1995).

Two gene products derive from *LIG3* which differ at their C-termini and are the result of alternative splicing. DNA ligase IIIα is ubiquitously expressed and contains a BRCT motif at its C-terminus, while DNA ligase IIIβ is only expressed in testes and ovaries (Mackey *et al.*, 1997, Perez-Jannotti *et al.*, 2001). Both isoforms contain at its N-terminus a zinc finger, homologous to the one found in poly-(ADP-ribose)-polymerase (Ikejima *et al.*, 1990, Mackey *et al.*, 1999) and by analogy it is thought to function as a nick sensor as well (Caldecott *et al.*, 1996). Expression analyses of DNA ligase IIIβ mRNA suggests that this isoform may be specifically involved in meiotic recombination (Mackey *et al.*, 1997).

DNA ligase IIIα interacts with the XRCC1 DNA repair protein *in vivo* and *in vitro* (Caldecott *et al.*, 1994, Caldecott *et al.*, 1995) and the interaction is mediated through its unique C-terminus where the BRCT-domains are located (Nash *et al.*, 1997, Taylor *et al.*, 1998a). XRCC1 also interacts with DNA polymerase β, which participates in BER (base excision repair) and cells lacking XRCC1 are deficient in short-patch BER (Caldecott *et al.*, 1996, Cappelli *et al.*, 1997). Furthermore, intact XRCC1 is required for normal cellular levels of DNA ligase III (Ljungquist *et al.*, 1994, Caldecott *et al.*, 1995). Consequently, DNA ligase IIIα is most likely to function in nuclear base-excision repair (Cappelli *et al.*, 1997). Generally, the base-excision repair pathway is responsible for repairing DNA damage that is induced by reactive oxygen species and small base adducts. After removal of the damaged base by a DNA glycosylase and an AP endonuclease, completion of BER can be divided into the short-patch pathway, where only one nucleotide is replaced, and the long-patch pathway, where displacement synthesis of 9-12 nucleotides occurs. DNA

ligase III/XRCC1 participate in short-patch BER (reviewed in Wood, 1996, Tomkinson *et al.*, 2001).

By 1976 a DNA ligase activity had first been described in rat liver mitochondria (Levin and Zimmerman, 1976). Recently, several proteins involved in BER have been purified from *X. laevis* mitochondria, amongst them a DNA ligase, that has been shown to be capable of repairing abasic sites in DNA (Pinz and Bogenhagen, 1998). The ligase activity of human mitochondria is encoded by the DNA ligase III gene and a mitochondrial targeting sequence (MTS) at its N-terminus is responsible for directing the enzyme to mitochondria. The same is true for the mitochondrial DNA ligase of rodents and *Xenopus laevis* (Lakshmipathy and Campbell, 2000, Perez-Jannotti *et al.*, 2001). Translation initiation from an upstream start codon produces the mitochondrial form whereas translation from the previously identified downstream start codon results in the major nuclear form (Lakshmipathy and Campbell, 1999, see also section 1.4). Mitochondrial DNA ligase function is essential for maintaining mtDNA integrity and stability (Lakshmipathy and Campbell, 2001). This suggests, together with the inability to detect any additional DNA ligase proteins in mitochondria (Lakshmipathy and Campbell, 2001, Perez-Jannotti *et al.*, 2001) that, in addition to repairing DNA damage created by reactive oxygen species, mitochondrial DNA ligase is involved in mtDNA replication. The mitochondrial ligase function does not depend on XRCC1 (Lakshmipathy and Campbell, 2000).
The completed genome sequences of the yeasts *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* do not reveal any obvious homologues of DNA ligase III.

#### 1.2.3.3 DNA ligase IV

The existence of another DNA ligase in humans, DNA ligase IV, was discovered by searching for additional proteins which contain a sequence homologous to conserved motif VI in human cDNA libraries (Wei *et al.*, 1995). This DNA ligase shares the catalytic core with DNA ligases I and III but differs by an extended C-terminus harbouring two BRCT domains. The region between those two domains binds to

XRCC4 (Grawunder *et al.*, 1997, Critchlow *et al.*, 1997, Grawunder *et al.*, 1998) and a crystal structure of the complex has recently been solved (Sibanda *et al.*, 2001). Together with DNA-dependent protein kinase (DNA-PK) and the Ku70/Ku80 heterodimer the DNA ligase IV-XRCC4 complex carries out non-homologous end joining (NHEJ), the main pathway in double-strand break repair and V(D)J recombination in higher organisms. XRCC4 seems to stabilise DNA ligase IV since null cells have greatly reduced levels of DNA ligase IV (Modesti *et al.*, 1999). Inactivation of the LIG4 gene in mice leads to lethality in late embryonic stages and cultured cells lacking DNA ligase IV are defective in V(D)J recombination and display an increased sensitivity to ionising radiation (Barnes *et al.*, 1998, Frank *et al.*, 1998, Gao *et al.*, 1998).

Homologues of DNA ligase IV have been identified in the yeasts *Saccharomyces cerevisiae* (Wilson *et al.*, 1997, Teo and Jackson, 1997, Schar *et al.*, 1997, Ramos *et al.*, 1998), *Candida albicans* (Andaluz *et al.*, 1996, Andaluz *et al.*, 1999) and *Schizosaccharomyces pombe* (Manolis *et al.*, 2001). A XRCC4 homologue, Lif1p, has also been identified in budding yeast (Herrmann *et al.*, 1998). As expected, yeast DNA ligase IV is an essential component of NHEJ since deletion of the respective genes from the chromosomes resulted in cells defective in this process. Yeast genes encoding DNA ligase IV are non-essential for growth and display only a slight sensitivity to various DNA damaging agents (Wilson *et al.*, 1997, Teo and Jackson, 1997, Schar *et al.*, 1997, Ramos *et al.*, 1998, Manolis *et al.*, 2001), which indicates that another pathway, most probably homologous recombination, is the major mechanism to repair DNA double strand breaks originating from ionising radiation in yeast.

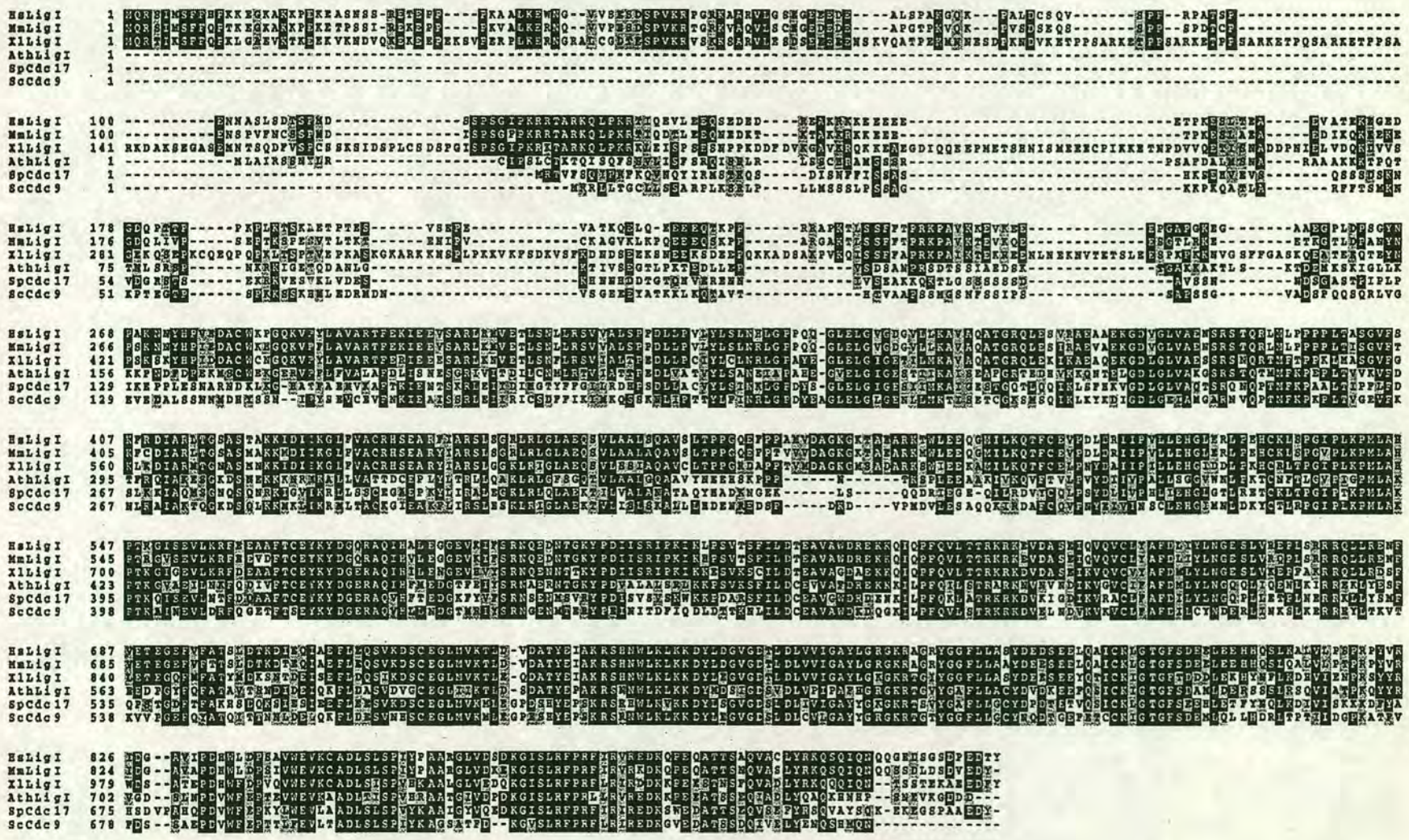
## 1.2.4 DNA ligase I

### 1.2.4.1 Introduction

By the mid-seventies, two distinct DNA ligases had been detected in mammalian cell extracts, only one of which, namely DNA ligase I, was induced upon cell proliferation. This enzyme is the major ligase activity in proliferating cells (Soderhall, 1975, Soderhall and Lindahl, 1975). The protein was first purified to apparent homogeneity from calf thymus (Teraoka and Tsukada, 1982; Tomkinson *et al.*, 1990).

A partial cDNA sequence of human DNA ligase I was obtained by screening a human cDNA library by hybridisation with oligonucleotides derived from a peptide sequence of bovine DNA ligase I. Subsequently, a full-length human cDNA was obtained by functional complementation of a *Saccharomyces cerevisiae cdc9* temperature-sensitive DNA ligase mutant (Barnes *et al.*, 1990, see also next section). To date, DNA ligase I homologues from human, mouse, *X. laevis*, *A. thaliana*, *S. pombe* and *S. cerevisiae* have been cloned and sequenced (Barnes *et al.*, 1990, Savini *et al.*, 1994, Lepetit *et al.*, 1996, Taylor *et al.*, 1998b, Barker *et al.*, 1987, Barker *et al.*, 1985). The size of these enzymes ranges from 180 kDa for *X. laevis*, 102 kDa for human down to 85 kDa for the budding yeast homologue. Figure 1.9 shows a sequence alignment of DNA ligase I homologues.

**Figure 1.9** (following page): Sequence alignment of DNA ligase I proteins from humans (HsLigI), mouse (MmLigI), *Xenopus laevis* (XlLigI), *Arabidopsis thaliana* (AthLigI), *Schizosaccharomyces pombe* (SpCdc17) and *Saccharomyces cerevisiae* (ScCdc9). The alignment was performed with ClustalX, using the default settings.



### 1.2.4.2 DNA ligase I in DNA replication

Several lines of evidence strongly imply a role for DNA ligase I in DNA replication. The budding yeast and fission yeast temperature-sensitive mutants *cdc9* and *cdc17*, respectively, that are deficient in DNA ligase I, arrest in early S-phase and accumulate low molecular weight DNA (Johnston and Nasmyth, 1978; Nasmyth, 1977; see also next section). Similarly, the human cell line 46BR, which contains two point mutations in different alleles of the *LIG1* gene, displays an abnormal joining of Okazaki fragments (Barnes *et al.*, 1992; Prigent *et al.*, 1994). Human DNA ligase I can complement the *Saccharomyces cerevisiae cdc9* mutant (Barnes *et al.*, 1990) and the defect in the joining of Okazaki fragments can be complemented by the addition of human DNA ligase I but not by DNA ligase III or T4 ligase (Mackenney *et al.*, 1997). Furthermore, in reconstitution of complete SV40 replication with purified replication factors, only mammalian DNA ligase I functions in the replication system whereas DNA ligase III cannot substitute and T4 ligase only functions inefficiently (Waga *et al.*, 1994). In summary, DNA ligase I is the enzyme responsible for the joining of Okazaki fragments that are generated during the DNA replication of the lagging strand.

The DNA ligase I gene is essential for cell viability. Both yeast temperature-sensitive mutants that are deficient in DNA ligase I are lethal at the restrictive temperature (Nasmyth, 1977; Johnston and Nasmyth, 1978) and gene knockout experiments with mouse embryonic stem cells – where both *LIG1* alleles had been inactivated – showed that the cells were only viable if a human DNA ligase I cDNA was ectopically expressed (Petrini *et al.*, 1995). However, a seemingly contrary report described that mouse embryos in which the last three exons had been deleted from both copies of the *LIG1* gene developed to midterm (Bentley *et al.*, 1996). It is possible that other ligases were partly substituting for DNA ligase I and/or low levels of undetectable ligase I protein was present.

### 1.2.4.3 Role of the N-terminal domain for DNA replication

Most studies with DNA ligase I have been carried out with the human enzyme. Human ligase I is composed of 919 amino acids with a molecular weight of 102 kDa, running on a SDS-gel at an apparent molecular weight of 125 kDa (Barnes *et al.*, 1990). It comprises a large C-terminal portion (amino acids 217–919) which can be sub-divided into the conserved non-catalytic domain and the T7-like catalytic core domains, and a small N-terminal domain (amino acids 1-216) which is hydrophilic and very sensitive to proteolysis. The N-terminal domain is not required for catalytic activity (Tomkinson *et al.*, 1990) but is indispensable for *in vivo* function. Mouse embryonic stem cells with a DNA ligase I null mutation are rescued by a full-length human DNA ligase I cDNA but not by an N-terminally truncated version (Petrini *et al.*, 1995). Interestingly, a human cDNA clone lacking the first 633 nucleotides, which corresponds roughly to the N-terminal domain, was able to complement the yeast *cdc9-7* mutant (Barnes *et al.*, 1990).

Studies with extracts of SV40-transformed 46BR cells, which displayed aberrant *in vitro* DNA replication, elucidated that full length DNA ligase I was 10-fold more efficient in complementation experiments than the catalytic domain on its own. This implies that the N-terminal part of DNA ligase I is required for efficient lagging strand replication (Mackenney *et al.*, 1997).

Several studies have been performed concerning the intracellular localisation of DNA ligase I throughout the cell cycle (Lasko *et al.*, 1990a, Li *et al.*, 1994, Montecucco *et al.*, 1995, Cardoso *et al.*, 1997). It has been found that in non-S-phase nuclei DNA ligase I is distributed throughout the whole nucleus excluding the nucleolus (Lasko *et al.*, 1990a). In mitosis, DNA ligase I is excluded from the condensed chromosomes and after re-formation of the nuclear envelope it is rapidly imported into the nucleus (Cardoso *et al.*, 1997). In S-phase nuclei DNA ligase I specifically localises at sub-nuclear sites of DNA synthesis (Montecucco *et al.*, 1995). Montecucco and co-workers identified a 13-amino acid stretch within the N-terminal domain of (human) DNA ligase I (amino acids 119-131) that serves as a nuclear localisation signal (NLS). They also found that the N-terminal amino acids preceding the NLS are necessary for the association of DNA ligase I with replication

factories (Montecucco *et al.*, 1995). In a later report, Cardoso *et al.* (1997) came to the conclusion that a bipartite sequence consisting of aa 1-28 and aa 111-179 were necessary and sufficient to direct DNA ligase I to sites of ongoing replication. The entire catalytic domain was completely dispensable for localisation to replication foci. DNA ligase I co-localised with PCNA and methyltransferase, which had previously been shown to be present at sites of DNA synthesis (Leonhardt *et al.*, 1992). Subsequently Montecucco and co-workers (1998) refined the mapping of the so-called replication factory targeting sequence (RFTS). The first 20 amino acids are necessary and sufficient for sub-nuclear targeting to replication foci *in vivo*.

DNA ligase I interacts with PCNA, a protein central to DNA replication, other DNA metabolic events and cell cycle progression (see section 1.1.3). Human PCNA was recovered as a protein specifically retained by a DNA ligase I affinity column and this interaction is mediated by the N-terminal 118 amino acids of DNA ligase I (Levin *et al.*, 1997). Intriguingly, in a later report it was shown that the first twenty amino acids of human DNA ligase I are sufficient for interaction with PCNA in GST-pull-down-assays – the same residues that are responsible for sub-nuclear targeting (Montecucco *et al.*, 1998) and that are conserved in various PCNA binding proteins (Warbrick, 1998, see section 1.1.3.4). Mutations of the conserved phenylalanine residues within the PCNA-binding motif abolish both targeting to replication foci as well as PCNA binding (Montecucco *et al.*, 1998). It is therefore tempting to speculate that DNA ligase I could be recruited to sites of DNA replication via an interaction with PCNA.

DNA ligase I and p21, which both contain the PCNA binding motif, inhibit *in vitro* PCNA-dependent DNA synthesis by pol δ (Levin *et al.*, 1997, Mossi *et al.*, 1998) which suggests that the former molecules compete with pol δ for binding to PCNA. Moreover, human DNA ligase I where the adjacent phenylalanine residues of the PCNA binding motif had been mutated, results in deficient joining of Okazaki fragments in 46BR.1G1 cell extracts whereas the wild type enzyme corrects the replication defect under those conditions (Levin *et al.*, 2000). Consistently, PCNA has been reported to increase ligation efficiency on a linear substrate and on a circular substrate in the presence of RF-C (as clamp-loader) (Tom *et al.*, 2001).

Taking into account the fact that DNA ligase I binds to PCNA via its N-terminal domain which is highly susceptible to proteolysis, it is understandable that the identification of this interaction is very much dependent on the experimental procedures. Loor *et al.* (1997), for example, failed to identify this interaction probably because of proteolytic degradation of the N-terminus of DNA ligase I.
In summary, evidence suggests that the observed interaction between DNA ligase I and PCNA is functionally relevant. However, the physiological significance of this association is still not answered conclusively.

The N-terminal portion of the protein also seems to be the target for post-translational modification; specific phosphorylation of serine residues by casein kinase II (CKII) within the N-domain modulates catalytic activity *in vitro*. Human ligase I expressed in *E.coli* is inactive. However, catalytic activity can be restored by preincubation with CKII (Prigent *et al.*, 1992). A residue in the N-terminal domain of human DNA ligase I, Ser66, which is part of a strong CKII consensus site, is phosphorylated in a cell-cycle dependent manner with it being phosphorylated in $G_2$ and mitosis, becoming dephosphorylated in early $G_1$ and the amount of phosphorylated form increasing during S-phase (Rossi *et al.*, 1999). Also, Cdk2-cyclin A is complexed with PCNA from HeLa nuclear extracts, this ternary complex could phosphorylate ligase I and PCNA seems to exert a stimulating effect on this activity (Koundrioukoff *et al.*, 2000).

### 1.2.4.4 DNA ligase I in DNA repair

In addition to its role in DNA replication, DNA ligase I also plays a role in DNA repair. The human 46BR cell line and the budding and fission yeast mutants are hypersensitive to a variety of DNA damaging agents such as UV light, alkylating agents and ionising radiation. Complementation of *cdc9* with human DNA ligase I cDNA corrects this phenotype (Barnes *et al.*, 1992; Johnston and Nasmyth, 1978; Nasmyth, 1977). However, it has been reported that the ligation activity in this case can be substituted for by ligase III or T4 ligase (Kubota *et al.*, 1996).

DNA ligase I is required for the final step in several DNA repair pathways. Nucleotide excision repair (NER) has been reconstituted with recombinant human proteins: in addition to a set of proteins specific for recognition and removal of the damaged site, the replication proteins DNA polymerase δ or ε, PCNA, RF-C and DNA ligase I were required for completion of DNA repair (Araujo *et al.*, 2000).

DNA ligase I has been found to interact with pol β, a DNA polymerase functioning in BER pathways. The two proteins were present in a complex capable of carrying out BER and co-immunoprecipitated from bovine testes nuclear extract (Prasad *et al.*, 1996). Subsequently, the interaction site within DNA ligase I was determined to lie in the N-terminal domain (Dimitriadis *et al.*, 1998). The precise residues required for the interaction are not known but mutations in the PCNA binding motif of human DNA ligase I did not affect PCNA binding suggesting that pol β and PCNA bind to distinct regions within the N-terminal domain. A functional PCNA binding motif is required in long-patch BER assays indicating that DNA ligase also acts in conjunction with PCNA in DNA repair pathways (Levin *et al.*, 2000)

### 1.2.4.5 Yeast DNA ligase I homologues

In a screen for temperature-sensitive mutants defective in cell cycle progression in *S. pombe*, mutants probably defective in the structural gene for DNA ligase, termed *cdc17*, were isolated (Nasmyth, 1977) (see also section 1.3). The cells grew normally at 25°C but, in contrast to wild type cells, were unable to form colonies at 35°C. Instead, the cells arrested in S-phase with a highly elongated phenotype (*cdc* phenotype). The mutants displayed four different phenotypes:

1) A lethality at the restrictive temperature (*cdc*$^-$) that is caused by a defect at the end of S phase - the cells underwent bulk DNA synthesis but the DNA was defective. 2) The mutants showed an abnormally high sensitivity to UV radiation, 3) a deficient joining of nascent DNA strands and 4) an abnormally low level of DNA ligase activity. Three different mutant alleles of the *cdc17* gene, arisen from independent mutageneses, have been isolated - *cdc17-K42*, which is the best-characterised allele, *cdc17-M75* and *cdc17-L16*. In 1979, Nasmyth described the partial purification of

Cdc17 and a new method of assaying DNA ligase. He showed that the *cdc*⁻ phenotype and the deficiency in DNA ligase activity in *cdc17*-K42 is due to a single mutation resulting in a thermolabile DNA ligase. Around the same time the *Saccharomyces cerevisiae* cell cycle mutant *cdc9* has been identified which showed similar properties to the *cdc17* mutants (Johnston and Nasmyth, 1978).

In the following years, both the *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* DNA ligase genes, *CDC9* and *cdc17*, respectively – homologues of mammalian DNA ligase I – were cloned by complementation of the *S. pombe cdc17* temperature-sensitive DNA ligase I mutant (Barker and Johnston, 1983; Johnston *et al.*, 1986) and subsequently sequenced (Barker *et al.*, 1985; Barker *et al.*, 1987).

*S. pombe* Cdc17 consists of 768 amino acids with a molecular weight of 86,581 Da and the *S. cerevisiae* counterpart comprises 755 amino acids and has a molecular weight of 84,828 Da. In contrast to *CDC9*, which does not have introns (Barker *et al.*, 1985), *cdc17* contains two small introns of 40 and 63 nucleotides, respectively, close to the 5' end (Barker *et al.*, 1987, see Appendix 1). The lack of introns in *CDC9* probably explains why *CDC9* is able to cross-complement *cdc17* mutations (Barker and Johnston, 1983) whilst the reverse complementation is not successful (Johnston *et al.*, 1986).

The yeast ligases contain the six boxes of homology common to DNA ligases and the active lysine residue of budding and fission yeast is at position 419 and 416, respectively. Like the human enzyme, the budding yeast homologue is an asymmetric monomer and has an N-terminal domain which is readily susceptible to endogenous and chemical proteolysis (Tomkinson *et al.*, 1990; Tomkinson *et al.*, 1992).

A striking difference between Cdc17 and Cdc9p is their mode of expression throughout the cell cycle. The *cdc17* transcript is present at constant levels whilst the *CDC9* m-RNA is periodically transcribed peaking at the G1/S boundary, and this also correlates with their specific activity of DNA ligation (White *et al.*, 1986). Since both enzymes fulfil the same function it could be considered surprising that they are expressed in such a different way, however, this may simply reflect the fact that the two yeasts differ substantially in the organisation of their cell cycles and that periodic expression in budding yeast does not serve any particular purpose.

Very recently another DNA ligase protein apart from Cdc17 and Lig4 was reported to be encoded by the *S. pombe* genome. This gene, termed *lpp1*$^{+}$ (for ligase paralogue protein, Stuart MacNeill, personal communication) encodes a ligase most closely related to DNA ligase I homologues. Its catalytic domains share around 40% sequence identity (~60% homology) with Cdc17 and human DNA ligase I and its conserved non-catalytic domain is ~30% identical (~50% homologous) with those of DNA ligase I proteins. The gene is not essential for viability and possible functions are currently under investigation (Stuart MacNeill, personal communication).

## 1.3 Introduction to the fission yeast *S. pombe*

### 1.3.1 Overview

The fission yeast *Schizosaccharomyces pombe* is a unicellular ascomycete fungus. *S. pombe* cells are rod-shaped, grow by apical extension and divide by medial fission. It is only distantly related to the budding yeast *Saccharomyces cerevisiae* and recent phylogenetic analyses suggest that they are only as related to each other as either of them is to animals (Sipiczki, 2000, Heckman *et al.*, 2001). Phylogenetically, *S. pombe* is classified as an Archiascomycete. Its genome size is 13.8 Mb, and the DNA is contained on three chromosomes of 5.7, 4.6 and 3.5 Mb. Sequencing of the genome is now complete and data can be accessed via the Sanger Centre website under http://www.sanger.ac.uk/Projects/S_pombe/.

*S. pombe* is an excellent model system for studying gene function and its haploid and diploid states of existence (see below) allow easy genetic manipulation.

### 1.3.2 *S. pombe* life cycle

When *S. pombe* cells are provided with an abundance of nutrients, they enter the mitotic cell cycle, which is described in more detail in the next section. When deprived of nutrients several developmental fates are possible. Essentially, *S. pombe* cells are of two mating-types, termed $h^+$ and $h^-$. If the culture only contains cells of one mating-type, they will exit the mitotic cell cycle and accumulate in stationery phase. If both mating-types are present, cells of opposite mating-type conjugate and form a zygote, which then undergoes meiosis and sporulation. A zygotic ascus is generated containing four spores, which lie dormant until the nutritional conditions improve. Then spores will germinate and the mitotic cell cycle is re-entered. However, if diploid zygotes are re-supplied with nutrients, prior to the initiation of meiosis, they enter the mitotic cell-cycle. The growth mode of diploid cells is similar to that of haploid cells except that the former are longer (20-25 μm versus 12-15 μm at cell division) and wider (4-5 μm versus 3-4 μm at cell division) than the latter. In the laboratory, the generation time of *S. pombe* varies between two to four hours,

depending on the media (complex or minimal) and growth temperature. Figure 1.10 illustrates the life cycle of *S. pombe* cells.

### 1.3.3 *S. pombe* cell cycle

The mode of growth of fission yeast cells by length extension allows the stage of a cell within the cell cycle to be determined by measuring its cell length (Mitchison, 1957). *S. pombe* possesses a typically eukaryotic cell cycle with distinct $G_1$, S, $G_2$ and M phases. M, $G_1$ and S phase each occupy ~10% of one cell cycle and $G_2$ accounts for the remaining 70%. In rapidly growing cells, $G_1$ and S phase are completed soon after mitosis but before cell division is complete. Thus, the smallest, new-born cells are in G2 (see Figure 1.11).

There are two major control points in the mitotic cell cycle of *S. pombe*: in late $G_1$ before the initiation of S phase and in late $G_2$ before the initiation of mitosis. The point in $G_1$ when cells become committed to the mitotic cell cycle is termed START (Hartwell *et al.*, 1974, Nurse and Bissett, 1981). Passage of START and entry into S phase are dependent on prior completion of mitosis and cells reaching a certain cell size (Nurse, 1975, Nurse and Thuriaux, 1977). The control point in late $G_2$ determines the time of entry into mitosis. Passage of this $G_2$/M control point require that the cells reach a critical size as well as completion of S phase and repair of any DNA damage that may have occurred. The presence of these checkpoints ensure that mitosis is not initiated until the faithful duplication of the genome and that DNA synthesis does not begin until mitosis is complete (control of the *S. pombe* cell cycle is reviewed in detail in MacNeill and Nurse, 1997).

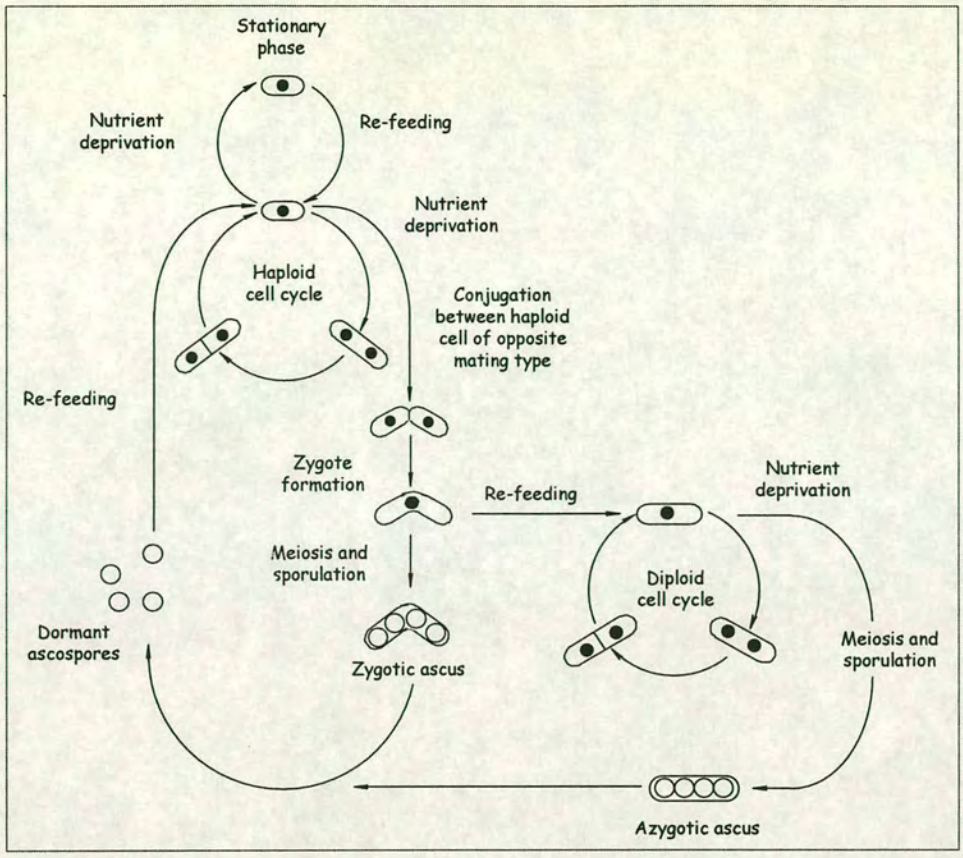



**Figure 1.10**: Life cycle of the fission yeast *S. pombe*. Top left: haploid mitotic cell cycle, centre and lower left: haploid cells mating to form a diploid zygote, followed by meiosis and sporulation resulting in a zygotic ascus. Lower right: re-entry of diploid cells into the mitotic cell cycle. Illustration kindly provided by Stuart MacNeill.

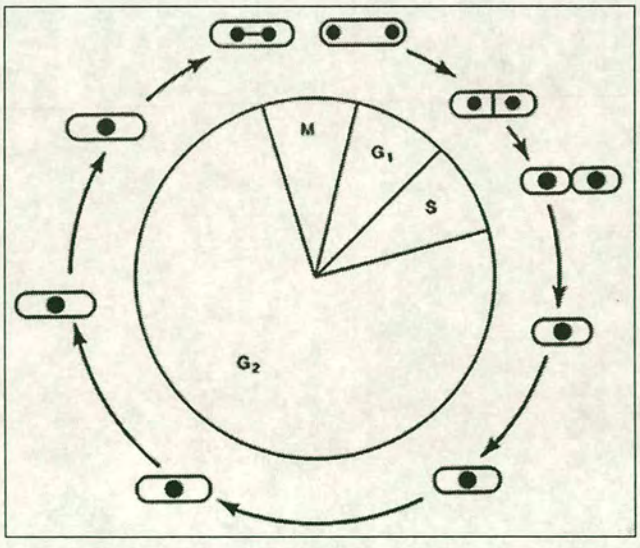



**Figure 1.11**: Schematic representation of the *S.pombe* cell cycle (picture taken from Alfa *et al.*, 1993). The relative duration of the cell cycle phases and morphological changes are indicated.

### 1.3.4 *cdc* mutants and S phase genes

Many conditional-lethal (i.e. temperature-sensitive or cold-sensitive) *S. pombe* mutants defective in cell cycle progression have been isolated in genetic screens. Cell cycle arrest in *S. pombe* typically leads to cell elongation where cells continue to grow but are unable to divide, due to checkpoint activation. Therefore, mutants can be identified by being highly elongated, termed the *cdc* (cell division cycle) phenotype. Twenty-five so-called *cdc* genes have been identified in the original screens for cell cycle defective mutants, which can be classified according to their arrest point. They include genes required for entry into the cell cycle, for DNA replication, for entry into mitosis and for septation (Nurse *et al.*, 1976, Nasmyth and Nurse, 1981). Eleven of the originally identified *cdc* mutants were shown to encode genes whose functions are required for S phase (reviewed in MacNeill and Nurse, 1997), among them *cdc17*, which is the subject of this work. Those genes are listed in Table 1.3 a). Table 1.3 b) gives an overview of additional genes whose products have subsequently been shown to be important during S phase in *S. pombe*. Finally, Figure 1.12 illustrates proteins involved in the *S. pombe* S phase.

**Table 1.3 a)**: *S. pombe cdc* genes, identified in the original screens for mutants defective in cell cycle progression (Nurse *et al.*, 1976, Nasmyth and Nurse, 1981), which have been shown to function in S phase. Different designations of the same gene are separated by a forward slash, and different gene products making up subunits of a functional complex are separated by commata.

| gene names | description |
|---|---|
| *cdc1* | DNA polymerase δ B-subunit |
| *cdc6/pol3* | DNA polymerase δ catalytic subunit |
| *cdc17* | DNA ligase |
| *cdc18* | regulator of S phase onset; Cdc6 homologue |
| *cdc19/nda1/mcm2* | Mcm2 |
| *cdc20* | DNA polymerase ε catalytic subunit |
| *cdc21/mcm4* | Mcm4 |
| *cdc22* | large subunit of ribonucleotide reductase |
| *cdc23* | Mcm10 |
| *cdc24* | required for completion of DNA replication; function unknown |
| *cdc27* | DNA polymerase δ C-subunit |

**Table 1.3 b)**: Further *S. pombe* S phase genes

| gene names | description |
|---|---|
| *cdm1* | DNA polymerase δ D-subunit |
| *cdt1* | regulator of S phase onset |
| *dna2* | Dna2 endonuclease/helicase |
| *hsk1, dfp1* | homologue of Cdc7 protein kinase |
| *mcm3, mcm5/nda4, mcm6/mis5, mcm7* | further MCM subunits (in addition to *mcm2/cdc19* and *mcm4/cdc21*) |
| *orp1/cdc30/orc1, orp2, orp3, orp4, orp5, orp6* | ORC subunits |
| *pcn1* | PCNA homologue |
| *pol1/cdc29/swi7* | DNA polymerase α catalytic subunit |
| *rad2* | FEN1 homologue |
| *rad4/cut5* | required for initiation of DNA replication and replication checkpoint |
| *rfc1, rfc2, rfc3, rfc4, rfc5* | RF-C subunits |
| *rpa1/rad11, rpa2, rpa3* | RPA subunits |
| *sna41* | Cdc45 homologue |

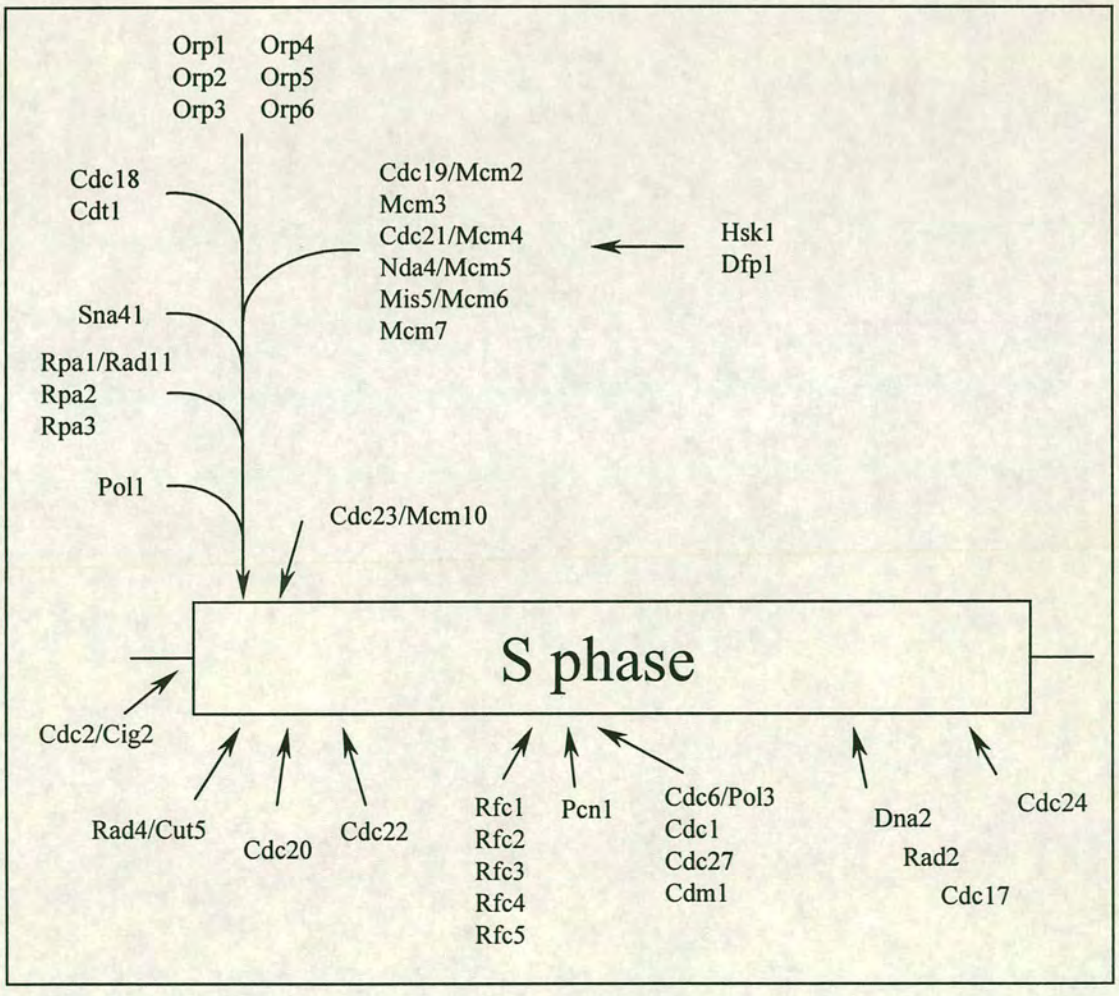



**Figure 1.12**: Schematic representation of *S. pombe* proteins involved in S phase.

## 1.4 Protein targeting to mitochondria

During the course of this work it became clear that *S. pombe* DNA ligase I, the subject of this study, is targeted to mitochondria – in addition to its nuclear localisation. The following section gives a brief introduction to mitochondria and subsequently an overview is given of mechanisms whereby targeting to mitochondria can be achieved.

### 1.4.1 Introduction

Mitochondria are organelles surrounded by a phospholipid bilayer membrane and are the sites of oxidative phosphorylation, whereby most of the ATP in eukaryotic cells is produced. They can be subdivided into the outer and inner membrane, the intermembrane space and the matrix. Mitochondria have their own circular DNA molecules, ranging in size from ~16.6 kb for human mitochondrial (mt) DNA, ~19.4 kb for the genome of *S. pombe* mitochondria to ~78 kb for *S. cerevisiae* mtDNA. The genomes encode between 8 and 13 proteins, the large and small subunits of rRNA that make up mitochondrial ribosomes, and between 22 and 25 tRNAs. All mitochondrially synthesised proteins are involved in electron transport or ATP synthesis. The genetic code used in mammalian and fungal mitochondria differs in most of the cases from the universal translation code; however, the standard code *is* used for protein-encoding genes in *S. pombe* mitochondria (Paquin *et al.*, 1997). Many proteins which are localised in mitochondria, such as DNA and RNA polymerases, metabolic enzymes and most proteins involved in oxidative phosphorylation, are encoded by nuclear genes and transcribed on cytosolic ribosomes. This raises the question as to how proteins are targeted to mitochondria. The following section gives an overview of common mechanisms that have evolved to ensure that proteins are targeted to mitochondria and other cellular destinations.

## 1.4.2 Targeting mechanisms

Nascent proteins, which are translated on cytosolic ribosomes, are usually targeted to organelles by uptake-targeting sequences specific for the respective location. For example, proteins destined for the endoplasmatic reticulum (ER) generally harbour an N-terminal signal sequence with a hydrophobic core whereas peroxisomal proteins have C-terminal tripeptides of the Ser-Lys-Leu motif. Nuclear localisation signals (NLS) are usually internal and consist of highly positively charged mono- or bipartite sequences. Mitochondrial targeting signals (MTS) are N-terminal sequences which form amphiphilic α-helices: one side of the helix being predominantly hydrophobic, whereas the other side is positively charged (von Heijne, 1986, Roise and Schatz, 1988). Apart from these common physico-chemical properties, no sequence similarities have been identified in mitochondrial presequences of different proteins. They are, however, capable of directing non-mitochondrial proteins into the organelle. The N-terminal MTS targets the so-called preprotein to mitochondria where it is recognised by the protein import machineries located in the mitochondrial membranes. These translocases are termed TOM and TIM (for translocase of the outer membrane and translocase of the inner membrane, respectively) and transport the target protein in energy-driven processes into the mitochondrial matrix. The detailed mechanisms of protein transport into mitochondria lie outside the scope of this introduction and are reviewed in Neupert, 1997 and Pfanner *et al.*, 1997. If the protein functions in mitochondrial compartments other than the matrix, further sorting processes come into play. In the matrix the targeting sequence of the translocated preprotein is cleaved off by a mitochondrial processing peptidase (MPP) and matrix chaperones aid folding of the protein into its active configuration.

How can multi-compartmentalised proteins that are encoded by a single gene contain more than one targeting information? The most common method by which isoforms of a particular protein localise to mitochondria as well as to another compartment, are produced, is to have two (or more) in-frame translation initiation sites in the 5' region of the gene. The region between the two start codons typically encodes an amphiphilic α-helix which serves as a mitochondrial targeting sequence. Two major pathways can be identified either of which leads to the presence of (generally) two

proteins, encoded by the same gene, being found in different sub-cellular compartments. Firstly, alternative transcription initiation from different promoters may produce two mRNAs, which differ at their 5' ends. The long transcript contains both AUG codons and default translation from the upstream AUG codon yields the mitochondrial preprotein. Translation from the short mRNA begins at the second AUG and results in the non-mitochondrial form. The relative abundance of the isoforms in this case are at least partly determined through the transcript levels. Alternatively, only a single transcript may be produced, but alternative translation initiation from in-frame AUG codons produce proteins with different N-termini. Parameters which can influence the choice of initiation codon include the secondary structure of the mRNA (Baim and Sherman, 1988) and the sequence context surrounding the start codons (Kozak, 1996). Irrespective of the strategy adopted, the longer protein usually contains the MTS and will therefore be targeted to mitochondria, whereas the shorter protein will remain in the cytosol unless it harbours additional sorting signals. Combinations of the above described mechanisms may yield more complex sub-cellular distributions.

One example of proteins which are targeted to their destinations in the above described manner is the gene products of *S. pombe uve1*$^+$. Fission yeast *uve1*$^+$ encodes a UV-damage endonuclease (UVDE), which functions in the UVDE-dependent repair pathway (UVER). *S. pombe* and *Neurospora crassa* both utilise this type of DNA repair system – closely related to BER (base-excision-repair) – for the removal of UV-induced DNA damage (Bowman *et al.*, 1994). The *uve1*$^+$ nucleotide sequence possesses three in-frame start codons at positions 1, 56 and 64 and the N-terminal 20 amino acids show characteristic MTS features. Translation initiation from the upstream AUG generates a protein which localises to mitochondria and usage of the second and/or third start codon yields nuclear UVDE. Repair kinetics of a UV-induced DNA damage (CPD, cyclobutane pyrimidine dimers) at a nuclear and mitochondrial locus indicate that UVER functions both in the nucleus and in mitochondria (Yasuhira and Yasui, 2000).

Higher eukaryotes have developed further strategies to allow production of multi-compartmentalised proteins from a single gene. For example, the products of the

human *UNG* gene are mitochondrial and nuclear uracil-DNA-glycosylase, designated UNG1 and UNG2, respectively. They are generated by a combination of alternative transcription initiation and alternative splicing (Nilsen *et al.*, 1997) and differ in their N-terminal residues. The N-terminus of UNG1 contains an MTS whereas N-terminal amino acids of UNG2 are part of a complex NLS (Otterlei *et al.*, 1998). Uracil-DNA-glycosylase initiates the first step in base-excision repair.

(Further examples of targeting mechanisms of a particular gene product to multiple cellular compartments are reviewed in Danpure, 1995.)

## 1.5 Aim of this work

This study initially set out to examine the functions of the non-catalytic domains of the *S. pombe* DNA ligase I protein Cdc17. In particular, to investigate by complementation analyses, whether the presence of the non-catalytic domains of Cdc17 were required for cell viability (Chapter 2). Previous work had shown that the N-terminal domain of Cdc17 was able to bind the *S. pombe* PCNA homologue Pcn1. Experiments that examine the relevance of this interaction for *in vivo* function of the ligase protein are described (Chapter 4). Prompted by recent publications and personal observations, the ability of $cdc17^{+}$ to encode a mitochondrial DNA ligase was also investigated (Chapter 3).

# 2 Genetics and structure-function analyses of *cdc17*

## 2.1 Introduction

As discussed in Chapter 1, in common with other DNA ligase I proteins, the *S. pombe* protein Cdc17 consists of three distinct domains. The C-terminal catalytic domains (Core) is highly conserved across all ATP-dependent ligases. The conserved non-catalytic domain (CNCD) is common to archaeal and eukaryotic DNA ligases and has no attributed function as yet. An N-terminal extension (NT) – unique to eukaryotic DNA ligase I homologues – shares little sequence similarity between species except a p21-like PCNA binding motif, which will be the subject of a later chapter (see Chapter 4).

There have been several reports on the effect of deleting the non-catalytic domains of DNA ligase proteins on their catalytic activity *in vitro* (Sekiguchi and Shuman, 1997 and see below). Clearly, the N-terminal domain is dispensable for catalytic activity in case of the human and budding yeast DNA ligase I proteins (Tomkinson *et al.*, 1990, Tomkinson *et al.*, 1992). In the absence of the conserved non-catalytic domain ligation activity was still detectable in *in vitro* assays albeit to a lower extent (Sekiguchi and Shuman, 1997). In human cells, the N-terminal domain was found to contain an NLS and a replication factory targeting sequence (RFTS) (Montecucco *et al.*, 1995 and 1998). However, the question whether this domain is essential for DNA ligase function has not been answered conclusively yet.

In this chapter the contribution of the non-catalytic domains of Cdc17, namely the conserved non-catalytic domain and the N-terminal domain, to the *in vivo* function of Cdc17 is investigated. Truncated versions of Cdc17 are tested for their ability to rescue temperature-sensitive *cdc17* strains and subsequently a *cdc17* deletion strain, which was constructed as part of this work.

In preparation for these functional complementation experiments the temperature-sensitive strains *cdc17-K42*, *cdc17-L16* and *cdc17-M75*, which were isolated as conditional lethal mutants defective in cell cycle progression at the restrictive temperature (Nasmyth, 1977), were characterised and their lesion sites determined.

## 2.2 Analysis of temperature-sensitive *cdc17* strains

### 2.2.1 Sequence characterisation of *cdc17*$^{ts}$ alleles

As a first step towards identifying domains within the Cdc17 protein that are required for cellular functions the lesion sites in the thermolabile proteins were determined. For this purpose, genomic DNA from the temperature-sensitive strains *cdc17-K42*, *cdc17-L16* and *cdc17-M75* was prepared and used as a template for PCR-amplification with primers flanking the *cdc17* locus. The alleles were cloned and sequenced (see Materials and Methods, section 6.10.2 for details).

In each case a single deviation from the wild-type DNA sequence was detected resulting in a single amino acid change. The results of the sequence analysis are presented in Table 2.1.

**Table 2.1**: Sequence analysis of mutant *cdc17* alleles

| allele | *K42* | *L16* | *M75* |
|---|---|---|---|
| nucleotide | 2216 | 3074 | 3371 |
| wild-type sequence | GAA | CCG | CCT |
| | Glu298 | Pro584 | Pro679 |
| mutant sequence | AAA | TCG | TCT |
| | Lys298 | Ser584 | Ser679 |

The numbering of nucleotides and amino acids is as in Appendix 1 and Appendix 2, respectively, and the relative positions of the mutations within the Cdc17 protein is depicted in Figure 2.1.

*cdc17-K42* carries a mutation within the conserved non-catalytic domain (CNCD) and the mutated sites of the ligase proteins of *cdc17-L16* and *cdc17-M75* lie within the catalytic domains. The positions of the lesions within Cdc17 are discussed in detail in section 2.6.

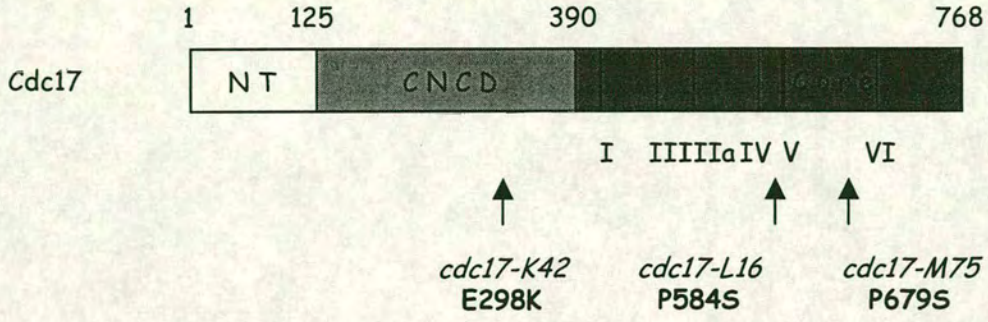



**Figure 2.1**: Positions of the mutation sites within the Cdc17 protein of the temperature-sensitive *cdc17* alleles. NT: N-terminal domain, CNCD: conserved non-catalytic domain, Core: catalytic domains. The roman numerals I – VI indicate the conserved motifs present in ATP-dependent DNA ligases. The numbers refer to amino acid positions.

## 2.2.2 Physiological characterisation

In addition to identifying the positions of the mutations, the growth properties of the temperature-sensitive strains at the permissive and restrictive temperature was compared. The cells were grown up overnight in YE at 25°C, split, diluted into fresh YE to an $OD_{600}$ of ~ 0.05 and grown at 25°C or 36°C. OD and cell number were monitored over a 10 hour period. The growth curves are shown in Figure 2.2.

In order to assess the functionality of the mutant proteins at permissive, intermediate and restrictive temperatures, cell length was measured after growth for 6 hours in YE at the respective temperatures (see Table 2.2).

**Table 2.2**: Cell lengths of *cdc17*$^{ts}$ strains at permissive, intermediate and restrictive temperatures

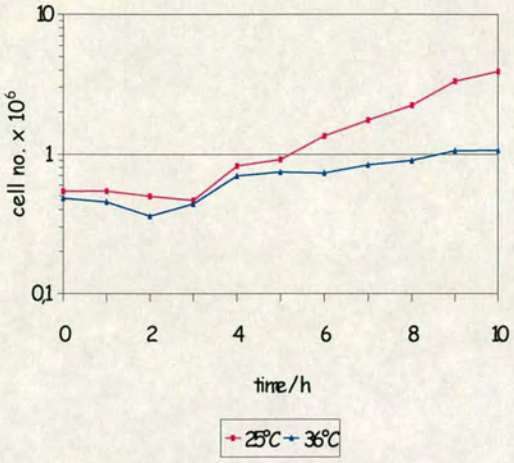
| strain | 25°C | 32°C | 36°C |
|---|---|---|---|
| *cdc17-K42* | (14.1 ± 1.7) µm | (23.5 ± 4.8) µm | (28.9 ± 4.7) µm |
| *cdc17-L16* | (13.1 ± 1.0) µm | (21.9 ± 4.6) µm | (29.8 ± 4.8) µm |
| *cdc17-M75* | (13.9 ± 1.2) µm | (14.7 ± 1.3) µm | (25.6 ± 5.1) µm |

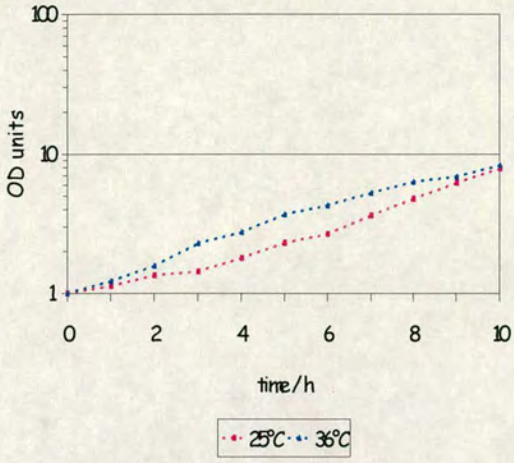
Strains *cdc17-K42* and *cdc17-L16* behave similarly. Although their cell mass increases to a similar extent at the permissive and restrictive temperature, cell number reached a plateau after around nine hours at 36°C, indicating a loss of viability. Consequently, the cells became highly elongated.

In contrast, with *cdc17-M75* the cell number increases similarly at the permissive and restrictive temperature over the examined time and the *cdc* phenotype manifests itself in a marked rise in optical density. Also, cell elongation in *cdc17-M75* is not as pronounced as in the other two strains. In summary, *cdc17-M75* shows a more subtle and less tight temperature-sensitive phenotype at 36°C.

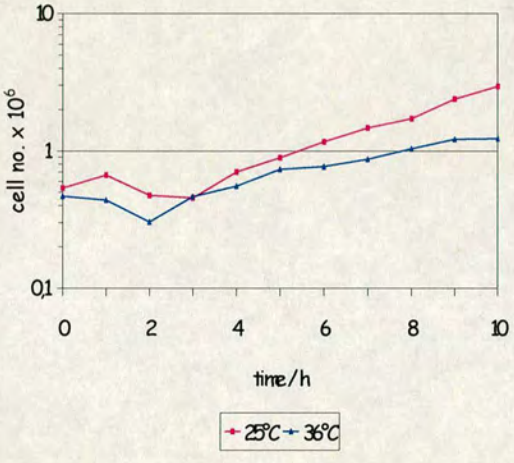
**Figure 2.2** (following page): Semi-logarithmic growth curves of a) *cdc17-K42*, b) *cdc17-L16* and c) *cdc17-M75*. Solid lines represent cell number (left graph) and dashed lines represent $OD_{600}$ (right graph). Growth properties at 25°C is indicated in pink and at 36°C in blue.

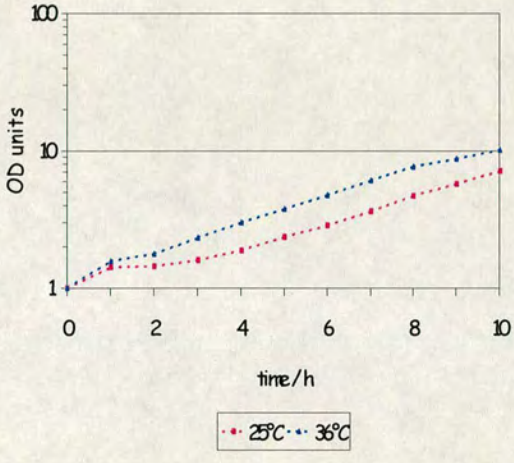
a) *cdc17-K42*



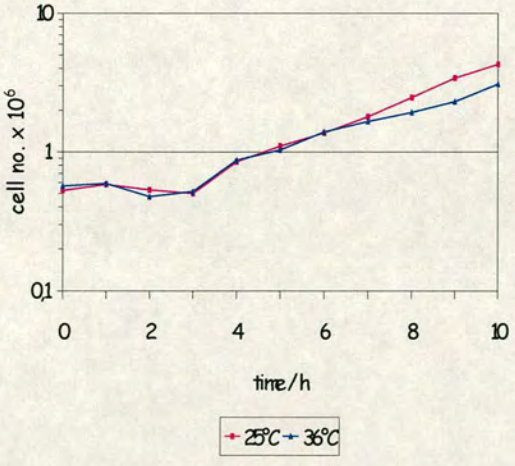



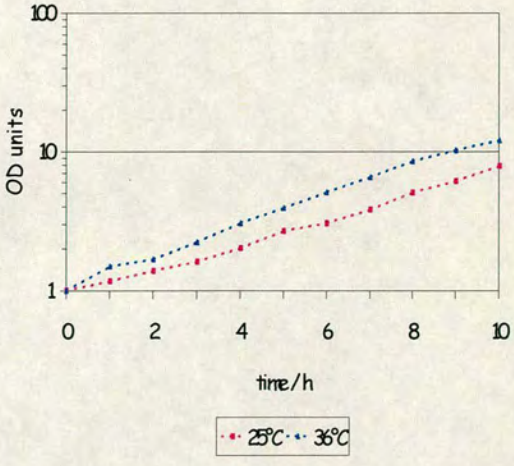

b) *cdc17-L16*




c) *cdc17-M75*



## 2.3 Cdc17 deletion strain

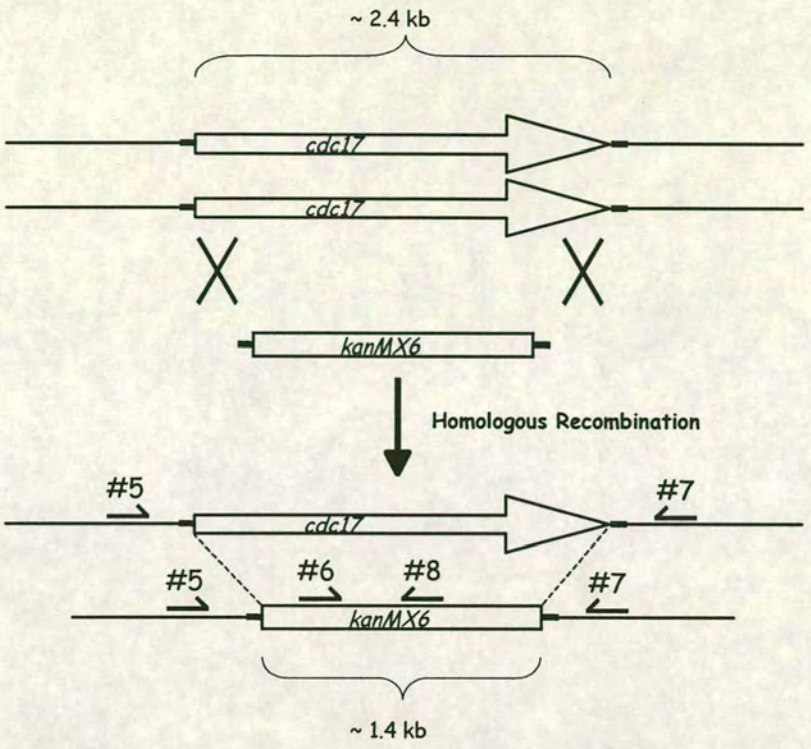
### 2.3.1 Introduction

In the previous section three temperature-sensitive *cdc17* strains were characterised. The extent of cell cycle inhibition at the restrictive temperature differed significantly among the strains. This indicates that those conditional lethal mutations are unlikely to be true null alleles and may retain some residual activity. In order to create a background that is completely depleted for the Cdc17 protein an *S. pombe* strain deleted for *cdc17* was engineered.

### 2.3.2 Construction

Since previous work (Nasmyth, 1977) has shown that *cdc17* is an essential gene, the deletion of *cdc17* was performed in a diploid strain. The complete gene was replaced with the *kanMX6* cassette, conferring resistance to the antibiotic G418. This gene deletion was carried out using the PCR-mediated gene targeting method described by Bähler *et al.*, 1998. Briefly, the *kanMX6* gene was PCR-amplified from the plasmid pFA6a-kanMX6 with primers that terminate in short stretches of homology (79 bp) to the genomic sequences immediately upstream and downstream of the *cdc17* coding sequence. 15 µg DNA was transformed into a diploid strain of genotype *leu1-32/leu1-32 ura4-D18/ura4-D18 his7-366/his7-366 ade6-M210/ade6-M216* $h^-/h^+$ and integrants were selected on G418 containing plates. Four stable transformants were recovered. In order to verify that the *kanMX6* module had integrated correctly at the *cdc17* locus, the transformants were screened by PCR. Genomic DNA from the four transformants and the corresponding wild-type strain was used as a template in PCR reactions with different combinations of primers. A schematic diagram of the construction of the *cdc17*$^+$*/cdc17*Δ*::kanMX6* diploid and the relative positions of the primers used for PCR are shown in Figure 2.3 a). Figure 2.3 b) shows a picture of an agarose gel with the PCR reaction products using genomic DNA of the four transformants and the wild-type strain as the template. For details see Materials and Methods, section 6.6.10.

a)




b)

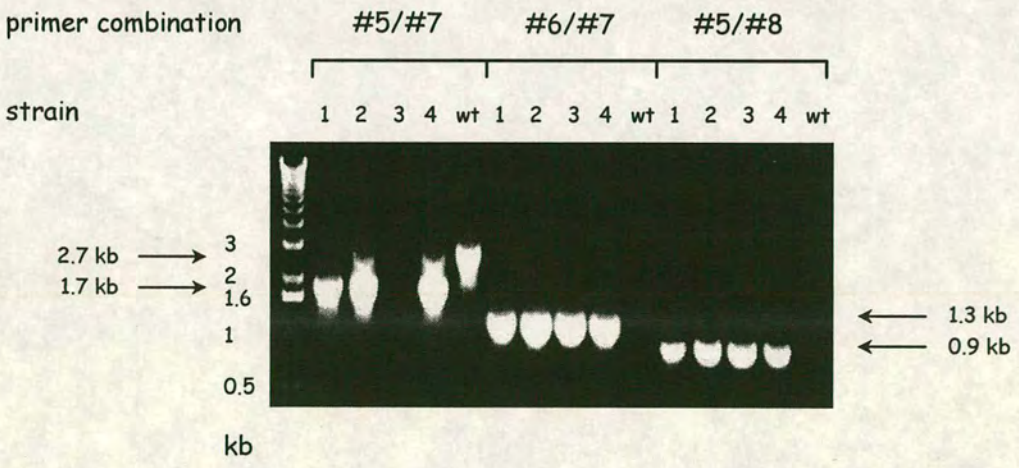



**Figure 2.3**: Replacement of one *cdc17* allele with the *kanMX6* cassette in a diploid strain. a) Schematic presentation of the PCR-based gene targeting method. The PCR-amplified *kanMX6* module with 5' and 3' ends homologous to sequences upstream and downstream of *cdc17* (indicated as thick lines) was introduced into a diploid strain (see text for genotype). Homologous recombination events should yield a heterozygous deletion diploid, as depicted in the lower half. The numbered arrows indicate primers which were used to amplify the region from the wild-type and putative deletion diploids using genomic DNA as a template. b) Agarose gel showing PCR reactions, where regions of genomic DNA from four putative heterozygous deletion diploids (named 1-4) and the corresponding wild-type was amplified with the indicated oligonucleotides (listed in Materials and Methods). In transformants 2 and 4 correct integration events had occurred.

### 2.3.3 Genetic analysis

Transformant 4 was used for further analyses. The heterozygous diploid of genotype *cdc17*$^{+}$*/cdc17::kanMX6 leu1-32/leu1-32 ura4-D18/ura4-D18 his7-366/his7-366 ade6-M210/ade6-M216 h*$^{-}$*/h*$^{+}$ was sporulated and 39 tetrads were dissected. Only in 10 cases did two colonies arise from the four isolated spores. In the other cases either only one spore resulted in a colony (20 out of 39) or all the four spores failed to produce a colony (9 out of 39). The reason for this is unclear. However, all the spores that were able to form colonies were G418-sensitive and adenine auxotrophs, i.e. haploids containing wild-type *cdc17*. Random spore analysis was also carried out and 96 meiotic products were all found to be G418-sensitive. Taken together the results from the spore germination experiments, no haploids deleted for *cdc17* could be recovered. This finding confirms the observation that *cdc17* is an essential gene – in agreement with Nasmyth, 1977.

## 2.4 Functional analyses

### 2.4.1 Introduction

In order to investigate Cdc17 function in several genetic backgrounds *cdc17* was cloned into the *S. pombe* expression vectors of the pREPX-series. pREP3X, pREP41X and pREP81X contain the thiamine regulatable *nmt* promoter with decreasing promoter strengths allowing expression levels which differ over five orders of magnitude (Maundrell, 1990, Maundrell, 1993, Basi *et al.*, 1993). The properties of the pREPX vectors are described in Materials and Methods, section 6.3.2). The full length *cdc17* gene was cloned into the *BamH*I site of the pREPX vectors.

## 2.4.2 Expression of *cdc17* in a wild-type background

In order to determine if high levels of Cdc17 exert any effect on wild-type cells *cdc17* was expressed at a variety of different levels: pREP3X, pREP3X-cdc17, pREP41X-cdc17 and pREP81X-cdc17 were transformed into the wild-type strain Sp5 (leu$^-$), grown to mid-log phase in EMM in the presence (promoter OFF) and absence (promoter ON) of thiamine and their cell lengths were measured. Table 2.3 gives an overview of the results.

**Table 2.3**: Cell lengths of wild-type strain Sp5 transformed with pREP3X, pREP3X-cdc17, pREP41X-cdc17 and pREP81X-cdc17

| plasmid | promoter OFF | promoter ON |
|---|---|---|
| pREP3X | (13.9 ± 1.6) µm | |
| pREP3X-cdc17 | (14.1 ± 1.2) µm | (22.3 ± 7.4) µm |
| pREP41X-cdc17 | (15.1 ± 1.4) µm | (15.1 ± 3.3) µm |
| pREP81X-cdc17 | (14.7 ± 1.6) µm | (14.6 ± 1.7) µm |

The highest possible expression of *cdc17*, i.e. from pREP3X in the absence of thiamine (promoter ON), leads to an elongation of wild-type cells of ~60% compared to control cells. Elongated *S. pombe* cells are symptomatic of a delay in cell cycle progression (see section 1.3) which, in this case, are a consequence of gross overproduction of Cdc17. An approximate ten-fold reduction in *cdc17* expression levels (expressed from the induced *nmt41* promoter) results in a cell length similar to wild-type. However, the increased standard deviation implies that a small proportion of cells are slightly elongated. This possibly reflects variations in plasmid copy number. It can be concluded that endogenous Cdc17 is produced to a lower extent than expression from pREP41X in the absence of thiamine (promoter ON). Expression of *cdc17* from the low strength *nmt81* promoter does not exert any visible effect in wild-type cells.

### 2.4.3 Rescue of *cdc17* mutant strains with *cdc17*$^{+}$

In order to estimate the minimal requirement for levels of functional Cdc17 protein in *S. pombe* cells, *cdc17* mutant strains were tested for rescue by various expression levels of *cdc17*. pREP81X-cdc17 was transformed into the temperature-sensitive strains *cdc17-K42* and *cdc17-M75*, grown to mid-log phase in the presence of thiamine, i.e. lowest possible expression level, and the cell lengths were determined. Table 2.4 summarises the measured cell lengths. Expression levels of *cdc17* from the repressed *nmt81* promoter are insufficient to rescue the *cdc* phenotype of *cdc17-K42*, since the cells are somewhat elongated at the restrictive temperature. In contrast, *cdc17-M75* can be rescued under the same conditions. In keeping with the observation that *cdc17-M75* is only mildly temperature-sensitive (see section 2.2.2), this may suggest that residual Cdc17-M75 activity contributes to the observed rescue.

**Table 2.4**: Cell lengths of the indicated temperature-sensitive *cdc17* strains transformed with pREP81X-cdc17, grown at the permissive temperature of 25°C or the restrictive temperature of 36°C in the presence of thiamine (promoter OFF)

| | 25°C | 36°C |
|---|---|---|
| *cdc17-K42*/pREP81X-cdc17 | (15.0 ± 1.5) μm | (18.5 ± 6.8) μm |
| *cdc17-M75*/pREP81X-cdc17 | (15.4 ± 2.3) μm | (14.5 ± 3.2) μm |

Additionally, meiotic products of *cdc17*$^{+}$/*cdc17*::kanMX6 transformed with either pREP41X-cdc17 or pREP81X-cdc17, which were G418-resistant and therefore deleted for *cdc17*, were grown up in the presence (promoter OFF) or absence (promoter ON) of thiamine at 32 °C and their cell lengths, which are shown in Table 2.5, were determined.

**Table 2.5**: Cell lengths of haploid *cdc17*Δ transformed with pREP41X-cdc17 or pREP81X-cdc17

| | promoter OFF | promoter ON |
|---|---|---|
| *cdc17*Δ/pREP41X-cdc17 | (15.3 ± 1.4) μm | (14.7 ± 3.0) μm |
| *cdc17*Δ/pREP81X-cdc17 | (18.8 ± 5.3) μm | (15.9 ± 1.8) μm |

As observed for *cdc17-K42*, expression of *cdc17* from the repressed *nmt81* promoter is not enough to fully rescue *cdc17*Δ, whereas an approximate ten-fold increase in expression level (i.e. expression from pREP41X, promoter OFF) does complement *cdc17*Δ.

Taken together, the above data indicate that the level of Cdc17 protein expressed from the repressed *nmt81* promoter is insufficient to fully rescue either *cdc17-K42* at 36°C or *cdc17*Δ.

### 2.4.4 Detection of Cdc17 protein in a wild-type background

Since cell lengths measurements indicate that neither *cdc17* expression from the induced *nmt*1 promoter nor from the repressed *nmt*81 promoter reflect endogenous levels of Cdc17 protein it was reasoned that expression from *nmt*41 yields Cdc17 quantities comparable to those found in wild-type cells. To confirm this, protein extracts were prepared from wild-type Sp5 cells transformed with either pREP41X or pREP41X-cdc17. Cells were grown in the absence and presence of thiamine and total protein was extracted from an equivalent amount of cells. Protein samples were run on an SDS polyacrylamide gel, transferred to a PVDF membrane and probed with anti-Cdc17-peptide antibody. Figure 2.4 shows a Western blot of the various cell extracts expressing Cdc17.

Wild-type cells expressing Cdc17 from the repressed *nmt41* promoter produce approximately 1.5 – 2 x the amount of Cdc17 compared to wild-type levels (Figure 2.4, lane 6 versus lane 5). Expression from the induced *nmt41* promoter leads to an 100 – 200-fold overproduction of the protein (compare lanes 3 and 4 to lane 6).

Supported by functional genetic analyses, it can be concluded that expression of *cdc17* from the repressed *nmt41* promoter is comparable to endogenous levels of Cdc17.

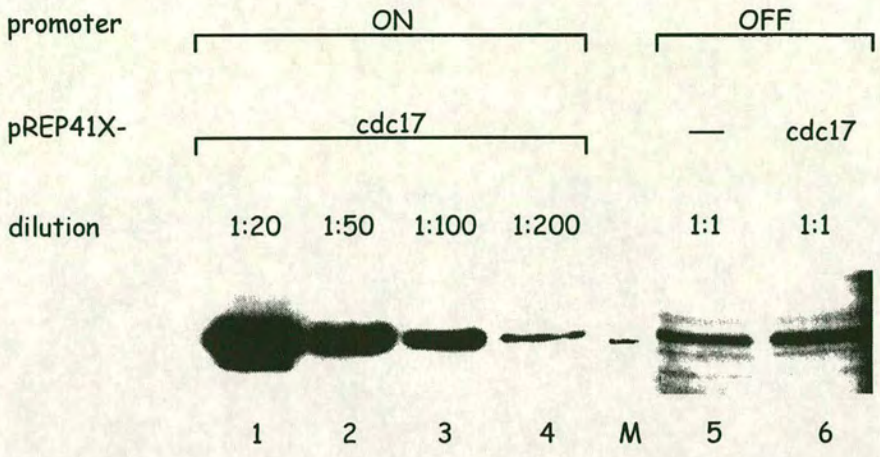



**Figure 2.4**: Serial dilutions of the protein extract from wild-type (Sp5) cells expressing Cdc17 in the absence of thiamine were made (lanes 1-4) in 1x SDS sample buffer and separated on an SDS polyacrylamide gel alongside protein extracts from Sp5/pREP41X (i.e. containing wild-type levels of Cdc17) and Sp5/pREP41X-*cdc17* grown in thiamine (lanes 5-6). The gel was blotted onto a PVDF membrane and probed with anti-Cdc17-peptide antibody used at a 1:5 dilution. Cdc17 protein was detected with the ECL system using the anti-rabbit secondary antibody at a 1:2000 dilution. lane M denotes a marker protein at 83 kDa.

## 2.5 Functional complementation experiments

### 2.5.1 Introduction

In order to determine which parts of the Cdc17 protein perform essential functions *in vivo*, several truncated constructs were engineered and tested for their ability to rescue temperature-sensitive strains as well as a deletion strain. The *cdc17* constructs were cloned into pREP41X since experiments described in the previous section have shown that production of Cdc17 from the repressed *nmt41* promoter reflects endogenous protein levels. The following section gives an overview of the constructs used before the results of the functional complementation experiments of the temperature-sensitive strains (section 2.5.3) and the deletion strain (section 2.5.4) are presented.

### 2.5.2 Cdc17 constructs

Several truncated or mutated forms of *cdc17* were constructed as described in Materials and Methods, section 6.3.3. *cdc17* genes encoding proteins lacking either the N-terminal domain or the N-terminal domain plus the conserved non-catalytic domain were cloned 3' to the *nmt* promoter into the *BamHI*-site of pREP41X. Full length *cdc17* served as a positive control and a construct with a mutation in the active site residue constituted a negative control. A description of the constructs used is given in Table 2.6 and Figure 2.5.

In order to confirm that the mutant Cdc17 proteins are expressed under the *nmt* promoter the pREP41X constructs were transformed into the wild-type Sp5 strain. Re-isolated transformants were grown up in EMM to mid-log phase and protein extracts from an equivalent amount of cells were prepared. The protein samples were separated on an SDS polyacrylamide gel and probed with anti-Cdc17-peptide antibody (1:200 dilution). A Western blot of wild-type cells expressing the various Cdc17 constructs is shown in Figure 2.6.

**Table 2.6**: pREP41X vectors, containing various mutant versions of *cdc17*, used for functional complementation experiments

| Plasmid name | Description of encoded protein |
|---|---|
| pREP41X-cdc17 | full length protein (amino acids 1-768) |
| pREP41X-cdc17-NΔ125 | Cdc17 lacking the N-terminal domain, but retaining conserved non-catalytic domain and catalytic core (amino acids 126-768) |
| pREP41X-cdc17-NΔ390 | Catalytic core only of Cdc17 (amino acids 391-768) |
| pREP41X-cdc17-K416A | Full length Cdc17 with a point mutation changing the active site residue lysine 416 to an alanine (negative control) |
| pREP41X | -------- |

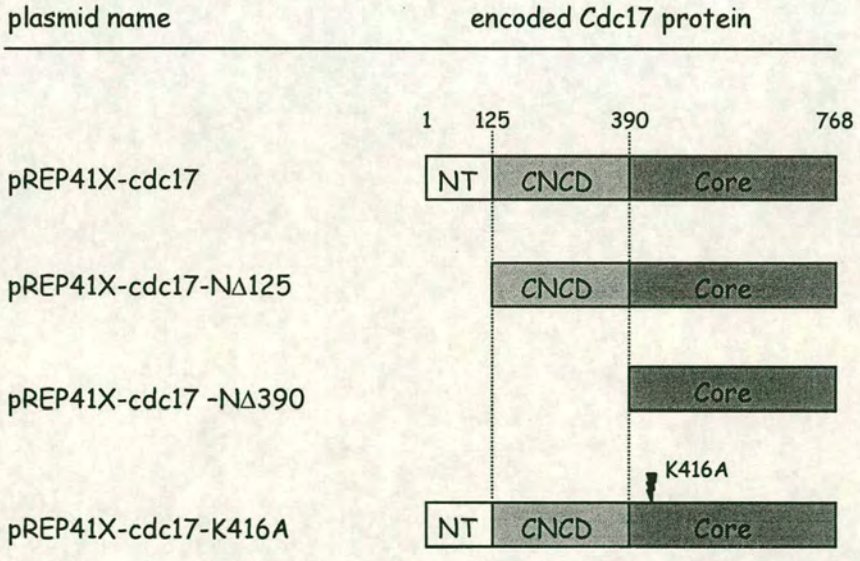



**Figure 2.5**: Schematic presentations of Cdc17 proteins, expressed from pREP41X, which are used in functional complementation experiments. Details of construct preparation are described in Materials and Methods.

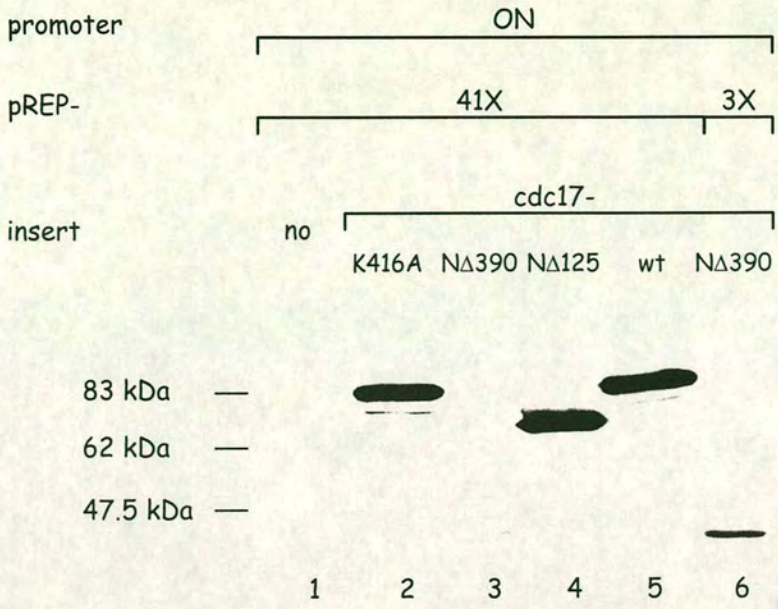



**Figure 2.6**: Protein extracts of wild-type Sp5 cells transformed with pREP-constructs were separated on an SDS polyacrylamide gel, which was blotted onto a PVDF membrane and probed with anti-Cdc17-peptide antibody used at a 1:200 dilution. This antibody is directed against the C-terminal 20 amino acids (aa 749-768) of Cdc17. Protein was detected with the ECL system using the anti-rabbit secondary antibody at a 1:5000 dilution. Lanes 1-6 contain an equivalent amount of protein extracts from the following cells: lane 1: Sp5/pREP41X, lane 2: Sp5/pREP41X-cdc17-K416A, lane 3: Sp5/pREP41X-cdc17-NΔ390, lane 4: Sp5/pREP41X-cdc17-NΔ125, lane 5: Sp5/pREP41X-cdc17, lane 6: Sp5/pREP3X-cdc17-NΔ390.

Full length Cdc17, the inactive ligase Cdc17-K416A and the Cdc17-NΔ125 protein, truncated for the N-terminal domain, are expressed to a similar extent (see lanes 5, 2 and 4) and are, as shown in section 3.4, 100 – 200-fold overproduced relative to endogenous Cdc17. In contrast, the level of Cdc17-NΔ390 protein is much lower than those of the above mentioned proteins (lane 3 versus lanes 2, 4 and 5). Even expression under the control of the strong *nmt1* promoter resulted in only marginallly increased Cdc17-NΔ390-levels (lane 6) indicating that the latter protein may not be very stable. However, Cdc17-Δ390 is present to a higher degree than wild-type levels of the Cdc17 protein, which are not detectable under the conditions used (lane 1).

## 2.5.3 Rescue of *cdc17*$^{ts}$ strains

### 2.5.3.1 Cdc17 proteins

The temperature-sensitive strains *cdc17-K42* and *cdc17-L16* were transformed with pREP41X harbouring various versions of *cdc17* (see Table 2.6 and Figure 2.5), plated onto EMM containing thiamine and incubated at 25°C until colonies appeared. (The growth characteristics and cell lengths of *cdc17-M75* at different temperatures (see section 2.2.2) imply that its temperature-sensitive phenotype is less tight than that of the other two strains. For this reason, *cdc17-M75* was not used for the complementation experiments.) Well-isolated transformants from each combination of *ts* strain and plasmid were grown up to mid-log phase in EMM with and without thiamine. Cell number was determined and 1000 cells per plate were plated twice onto EMM with and without thiamine. Plates were incubated at 25°C and 36°C for five days.

Irrespective of the plasmids present in the two *ts* strains the cells were viable at the permissive temperature of 25°C. Consequently, the *cdc17* constructs do not exert any dominant negative phenotype at this temperature. However, at the restrictive temperature of 36°C the ability of the different Cdc17 proteins to rescue varies considerably among the alleles. The results of the rescue experiments are summarised in Table 2.7.

**Table 2.7**: Ability of mutant Cdc17 proteins to rescue the temperature-sensitive strains *cdc17-K42* and *cdc17-L16* at 36°C. Growth was scored subjectively on a scale of +++, which indicates excellent rescue similar to wild type, to --- for no rescue. The number of colonies as well as their phenotype, judged by microscopical examination (see footnotes), were taken into consideration.

| Strain | *cdc17-K42* | | *cdc17-L16* | |
|---|---|---|---|---|
| Promoter | OFF | ON | OFF | ON |
| pREP41X-cdc17 | +++ | +++ | +++ | +++ |
| pREP41X-cdc17-NΔ125 | ---[2] | ++[1, 2] | +++[1] | +++ |
| pREP41X-cdc17-NΔ390 | --- | --- | --- | --- |
| pREP41X-cdc17-K416A | --- | --- | --- | --- |
| pREP41X | --- | --- | --- | --- |

[1] cells are slightly elongated [2] microcolonies of ~16-500 small, round, swollen cells

On the plates incubated at the restrictive temperature there was a wide range of colony sizes present. This probably reflects a variation of expression levels in individual cells. The full length Cdc17 protein was able to rescue both *ts* alleles irrespective of the promoter being induced or repressed. The major difference between *cdc17-K42* and *cdc17-L16* is their capacity to be rescued by Cdc17-NΔ125. Whereas Cdc17 lacking the N-terminal domain rescues *cdc17-L16* quite well at low levels it cannot rescue *cdc17-K42*. Overproduction of Cdc17-NΔ125 partly restores viability of *cdc17-K42* and leads to full complementation of *cdc17-L16*. Interestingly, microscopic examination of *cdc17-K42* expressing low levels of Cdc17-NΔ125 revealed microcolonies of ≤500 cells, which were small, round and swollen. This phenotype is atypical of cells with defects in S phase suggesting that Cdc17 may carry out other non-S-phase-specific-functions, which are defective in *cdc17-K42* but not in *cdc17-L16*.

A Cdc17 protein consisting only of the catalytic domains, Cdc17-NΔ390, was unable to rescue either temperature-sensitive strain. Since immunoblotting analyses showed that this protein is expressed at much lower levels, its inability to complement the *ts* strains may be the result of insufficient protein present in the cell. In order to clarify

if higher levels of the core-protein Cdc17-NΔ390 rescued *cdc17*$^{ts}$ strains the corresponding gene portion was cloned under the control of the stronger *nmt1* promoter. Plasmid pREP3X-cdc17-NΔ390 was transformed into Sp5 cells and total protein extract was prepared from mid-log phase cells. The extracts were separated on an SDS polyacrylamide gel and a Western blot (see section 2.5.2, Figure 2.6, lane 6) illustrates the relative amount of protein compared to expression under the *nmt41* promoter. Expression of *cdc17*-NΔ390 from *nmt1* is several fold higher than from *nmt41*. pREP3X-cdc17-NΔ390 was transformed into *cdc17-K42* and c*dc17-L16*. Isolated transformants were streaked out on EMM alongside various controls and incubated at 25°C and 36.5°C.

Figure 2.7 shows pictures of the plates after incubation for 3 days at the respective temperatures. Production of Cdc17-NΔ390 at the highest possible level was unable to rescue either *cdc17-K42* or c*dc17-L16*. Considering that this level is still approximately 10 x below wild-type-level it is not possible to conclude that the core Cdc17 protein would be able to rescue at comparable levels. It was noted, however, that *ts* strains transformed with pREP3X-cdc17-NΔ390 and grown in the absence of thiamine (i.e. promoter ON) at the permissive temperature showed a pronounced increase in generation time. Thus, overproduction of the catalytic domains of Cdc17 results in a slight dominant negative phenotype. This growth inhibition is not due to a cell-cycle delay since the cells are not elongated.

The rescue experiments of *cdc17-K42* and c*dc17-L16* by truncated Cdc17 proteins demonstrate that the C-terminal catalytic domains alone are not sufficient to functionally complement strains temperature-sensitive for Cdc17. Consequently, the conserved non-catalytic domain of Cdc17 appears to perform essential functions.

However, the importance of the N-terminal domain for Cdc17 function cannot be conclusively assessed, since a protein lacking this domain was unable to rescue *cdc17-K42,* but did complement *cdc17-L16*. Clearly, the presence of the respective *ts* proteins in the mutant cells at the restrictive temperature contributes to the varying ability of truncated ligase proteins to restore function.

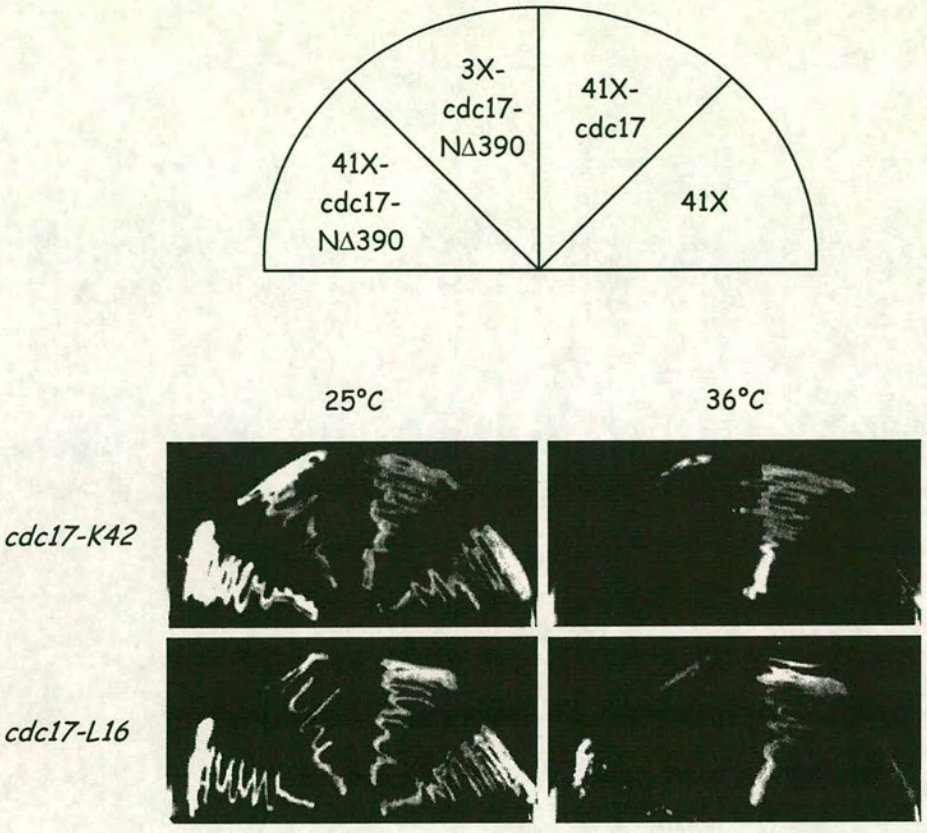



**Figure 2.7**: The temperature sensitive-strains *cdc17-K42* and *cdc17-L16* were transformed with the pREP-plasmids indicated in the top panel. Well-isolated transformants were streaked out on EMM and incubated at 25°C and 36°C for 3 days.

### 2.5.3.2 Other DNA ligase proteins

In addition to Cdc17 truncations other DNA ligase proteins were also tested for their ability to complement *cdc17*$^{ts}$ strains. The catalytic core domains of the putative *S. pombe* DNA ligase Lpp1 as well as the DNA ligase of the archaeon *Sulfolobus sulfataricus*, expressed from pREP3X and pREP41X, were examined regarding their ability to sustain growth of *cdc17-K42* and c*dc17-L16* (plasmids were provided by Stuart MacNeill). However, no rescuing activity could be detected.

In order to investigate if possible Lpp1 functions could partially overlap with those of Cdc17 the effect of absence of *lpp1* on *cdc17*$^{ts}$ strains was examined. A haploid strain deleted for *lpp1* (provided by Stuart MacNeill) was crossed to *cdc17-K42*, c*dc17-L16* and c*dc17-M75,* the resulting diploids were sporulated and the meiotic products inspected. In each case *lpp1*Δ *cdc17*$^{ts}$ double mutants were recovered with a percentage between 23-30%. This indicates that *lpp1* appears to function, if at all, independently of *cdc17*.

## 2.5.4 Rescue of *cdc17*Δ

### 2.5.4.1 Introduction

Functional complementation experiments of *cdc17*$^{ts}$ strains by an N-terminal Cdc17 truncation, described in the previous section, could not provide an unambiguous answer to the role of the N-terminal domain for *in vivo* function, possibly due to the varying contribution of the different thermolabile proteins. In order to examine the ability of truncated Cdc17 proteins *per se* to sustain growth of cells, rescue experiments were performed in a *cdc17*Δ strain.

### 2.5.4.2 Experimental Strategy - Plasmid Shuffle System

A common strategy to test essential protein functions in a deletion strain would be to introduce the expression constructs in a heterozygous deletion diploid and examine the cells carrying the deletion allele following sporulation. The *kanMX6* gene replacing one *cdc17* allele in the constructed *cdc17*Δ strain, and which confers resistance to the antibiotic G418, was found to be unsuitable as a selectable marker in minimal media, since wild-type cells (which do not have a copy of *kanMX6*) were able to grow on EMM plates containing G418.

Instead of selecting for *cdc17*Δ cells carrying the expression constructs in succession, it was decided to use a plasmid shuffle system. As a basis for the plasmid shuffle, a haploid *cdc17*Δ strain was constructed which is kept alive by a wild-type copy of the gene on the pUR19 plasmid, carrying the *ura4* gene which complements the uracil auxotrophy of the strain (genotype: *cdc17*Δ::*kanMX6 leu1-32 ura4-D18 his 7-366 ade6-M21X* pUR19-cdc17). The various pREP41X-cdc17 constructs ($leu^+$) were introduced into this strain, thereby complementing its leucine auxotrophy. Cells carrying both plasmids were plated out on 5-FOA (5-fluoroorotic acid), which selects against $ura^+$ cells. Therefore, the arising colonies should only contain pREP41X plasmids and the ability of truncated or mutated *cdc17* versions to rescue the deletion can be assessed. Details of the experimental procedure are described in Materials and Methods, section 6.6.11.

### 2.5.4.3 Complementation experiments of *cdc17*Δ

The pREP41X-vectors carrying mutant *cdc17* versions listed in Table 2.6 were transformed into *cdc17*Δ/pUR19-cdc17 ($ura^+$) and plated on media selecting for both plasmids either with or without thiamine. After isolation of transformants cells were grown up in EMM with and without thiamine to mid-log phase. Cell number was determined and 1000 cells were plated onto EMM with supplements, uracil, 5-FOA and either with or without thiamine. Plates were incubated at 32 °C for seven days.

The results of the rescue experiments are summarised below and in Table 2.8. Figure 2.8 shows photographs of most of the plates.

**Table 2.8**: Rescue of *cdc17*Δ by mutated Cdc17 proteins. The experiment was carried out as described in the previous section and Materials and Methods. + indicates excellent rescue and - denotes no rescue

| Promoter | OFF (+ thiamine) | ON (- thiamine) |
|---|---|---|
| pREP41X-cdc17 | + | + |
| pREP41X-cdc17-NΔ125 | -- | + |
| pREP41X-cdc17-NΔ390 | -- | -- |
| pREP41X-cdc17-K416A | -- | -- |
| pREP41X | -- | -- |

The full length Cdc17 protein is able to rescue *cdc17*Δ irrespective of the promoter being turned off or on, resulting in endogenous protein levels or 100 – 200-fold overproduction, respectively (Figure 2.8 a)). Cdc17-NΔ125, the protein truncated for the N-terminal domain, is capable of complementing *cdc17* deletion cells, but only when overexpressed (promoter ON). Expression of the Cdc17-NΔ125-truncation at levels comparable to those of the full length enzyme in wild-type cells (promoter OFF) does not give rise to any viable colonies, indicating that the N-terminal domain does perform important functions *in vivo* (Figure 2.8 b)). The Cdc17-NΔ390 protein, consisting only of the C-terminal catalytic domains, is not able to rescue under either condition (Figure 2.8 c)). Considering that, when the *nmt* promoter is induced, only a fraction of the amount of full length or N-terminally truncated protein is produced (see Figure 2.6), it cannot be completely ruled out that levels corresponding to full-length or the N-terminally truncated Cdc17-NΔ125 protein in the induced state were able to rescue. As expected, *cdc17*Δ cells are inviable in the absence of any Cdc17 protein (Figure 2.8 d)).

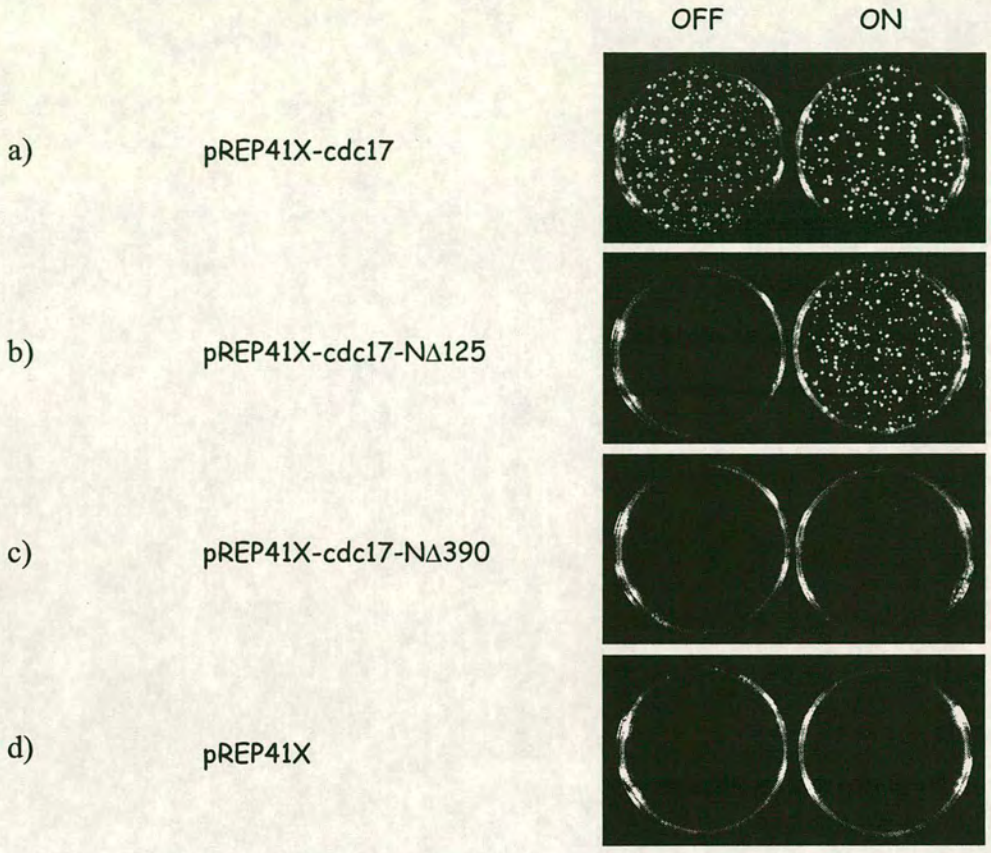



**Figure 2.8**: Rescue of *cdc17*Δ by the indicated *cdc17* constructs. Experimental details are described in the previous sections and Materials and Methods, section 6.6.11. Plates on the left side contain thiamine, which represses expression from the *nmt* promoter and plates on the right lack thiamine which induces expression of *cdc17* constructs under the *nmt* promoter.

## 2.6 Discussion

In the previous sections strains temperature-sensitive for *cdc17* have been characterised and examined for rescue with truncated Cdc17 proteins.

In rescue experiments, the *cdc17-K42* strain behaves most like the deletion strain, in that it is not complemented by a Cdc17 protein deleted for the N-terminal domain, Cdc17-NΔ125, when produced at levels comparable to wild-type. *cdc17-K42* carries a mutation within the conserved non-catalytic domain (CNCD). The negatively charged glutamic acid residue 298 is replaced with a positively charged lysine residue. Alignment of DNA ligase I homologues shows that Glu298 constitutes an invariant residue (see Figure 2.9 left) in a highly conserved area between eukaryotic and archaeal proteins. This degree of conservation suggests that essential functions are intrinsic to this domain. The mutated site of the ligase protein of *cdc17-L16* lies within the catalytic domains. Pro584, which is conserved between yeast (see Figure 2.9, bottom) and is situated between conserved motifs IV and V, is replaced by a serine residue. Unlike *cdc17-K42* and *cdc17*Δ, *cdc17-L16* could be complemented by Cdc17-NΔ125, expressed at levels similar to wild-type.

The above described lesions could manifest itself in several, possibly different, ways. Firstly, the protein levels after incubation at the restrictive temperature could vary due to differences in enzyme stability. Preliminary immunoblotting analyses of protein extracts from all three temperature-sensitive strains could not detect any significant decrease in mutant Cdc17 levels after growth for six hours at the restrictive temperature (data not shown). It can, however, not be ruled out that the enzyme stability does decrease, since the protein amounts were just above detection levels.

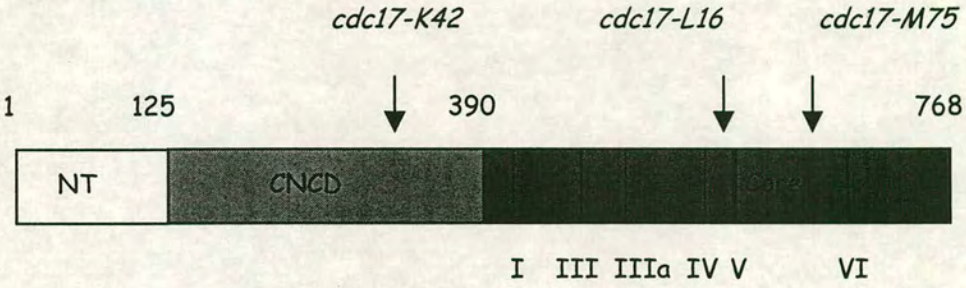
Secondly, *cdc17-K42* and *cdc17-L16*, which were employed for the rescue experiments, possibly show a varying degree of residual ligase activity: whereas Cdc17-L16 activity may just be below the threshold required for normal cell cycle progression and can therefore be complemented by low levels of a protein with intact catalytic domains, the thermolabile Cdc17-K42 activity may be much lower and therefore cannot be rescued. Since the N-terminal domain of Cdc17 does perform important functions, as judged by its presence being necessary for rescue of *cdc17*Δ

at low levels comparable to endogenous conditions, the N-terminal domain of Cdc17-K42, carrying out possible targeting functions, could direct the protein to sites of nicked DNA, but through the mutation in the conserved non-catalytic domain the enzyme is rendered deficient to complete its task. Thus, Cdc17-K42 may compete with the truncated Cdc17-NΔ125.

The mutation in *cdc17-L16* is situated between motifs IV and V, which are common to nucleotidyltransferases. The T7 crystal structure reveals that those residues are situated at the base of the cleft between the two catalytic domains and an optimal structural configuration is likely to be crucial for efficient catalytic mechanism and/or affinity for DNA, factors which may be reduced in this mutant.

In contrast, the altered residue in *cdc17-M75*, Pro679, which is converted to a serine, is unique to the fission yeast protein and falls within an amino acid stretch of comparably little sequence homology (see Figure 2.9, right). Consistent with this, the temperature-sensitivity of *cdc17-M75* can be considered mild: no significant difference in growth behaviour of *cdc17-M75*, grown for 10 hours at either the permissive or restrictive temperature, was seen. Instead, the observed cell elongation correlated with an increase in cell mass. Ligation assays with crude extracts or partially purified enzyme of *cdc17-M75* failed to detect any thermolability of the mutant protein *in vitro* (Nasmyth, 1979). It seems that the mutation in the *cdc17-M75* allele gives rise to only a slightly deficient protein, possibly due to minor structural changes which reduce enzyme kinetics.

**Figure 2.9** (see following page): The mutation sites of the three temperature-sensitive alleles of cdc17 within the protein are indicated. Sequence alignment of proteins from several species around the lesion sites of *cdc17-K42*, *cdc17-L16* and *cdc17-M75* are shown. Species names are as follows: *Sp: Schizosaccharomyces pombe, Sc: Saccharomyces cerevisiae, Hs: Homo sapiens, Xl: Xenopus laevis, Da: Desulfolobus ambivalens* (archaeon). The altered amino acids of the corresponding mutant proteins are shown above the alignment. Invariant residues are marked with "*", highly conserved residues with ":" and conserved residues with "". The sequence alignment was performed using ClustalX.




**K42**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | K | | | | | | | | |
| | . | | | * | : | : | : | . | * | | : | |
| *Sp* | C | E | G | A | **E** | P | K | Y | L | I | R | A | L |
| *Sc* | C | K | G | I | **E** | A | K | F | L | I | R | S | L |
| *Hs* | C | R | H | S | **E** | A | R | F | I | A | R | S | L |
| *Xl* | C | R | H | S | **E** | A | R | Y | I | A | R | S | L |
| *Da* | A | S | P | L | **E** | A | K | Y | L | V | R | F | V |

**M75**

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | S | | | | * | * |
| *Sp* | F | Y | A | H | S | D | V | **P** | A | H | Q | P | D |
| *Sc* | T | F | V | F | D | S | S | - | - | A | E | P | D |
| *Hs* | Y | V | R | I | D | G | A | - | - | V | I | P | D |
| *Xl* | Y | Y | R | W | D | S | A | - | - | T | E | P | D |
| *Da* | - | - | R | V | V | S | T | - | - | M | V | P | D |

**L16**

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | S | | | | | | |
| | : | . | | * | | | | . | | * | | | * |
| *Sp* | D | S | H | Y | E | **P** | S | K | R | S | R | H | W |
| *Sc* | E | S | H | Y | E | **P** | S | K | R | S | R | N | W |
| *Hs* | D | A | T | Y | E | **I** | A | K | R | S | H | N | W |
| *Xl* | D | A | T | Y | E | **I** | A | K | R | S | H | N | W |
| *Da* | D | A | I | Y | Q | **A** | G | S | R | G | W | L | W |

## 2.7 Summary

In this chapter the mutation sites of the three existing temperature-sensitive alleles of *cdc17* have been determined. *cdc17-K42* and *cdc17-L16* have been tested for functional complementation by Cdc17 proteins deleted for the N-terminal domain or only consisting of the catalytic domains. Additionally, a strain deleted for *cdc17* was constructed and tested for rescue by truncated proteins. *cdc17-K42*, which bears a point mutation in the middle conserved non-catalytic domain, as well as *cdc17*Δ was not complemented by an N-terminally truncated Cdc17 protein, when expressed at levels similar to wild-type. In contrast, this protein was able to rescue *cdc17-L16*, harbouring a mutation in the C-terminal catalytic core. Moderate overexpression rescued all three strains. The catalytic core of Cdc17 was unable to rescue either strain irrespective of the expression level. By genetic and biochemical methods it was established that endogenous levels of Cdc17 are comparable to expression from the repressed *nmt* promoter in pREP41X.

It is concluded that the conserved non-catalytic domain of Cdc17 performs essential functions and that the N-terminal domain contributes to proper functioning of the enzyme *in vivo*.

# 3. Nuclear and mitochondrial forms of Cdc17

## 3.1 Introduction

The N-terminal domain of DNA ligase I is poorly conserved from species to species, raising the question of how it contributes to DNA ligase function. Although the N-terminal domain of human DNA ligase I is dispensable for catalytic activity *in vitro* (Tomkinson *et al.*, 1990), it is required for *in vivo* function (Petrini *et al.*, 1995). Studies on human DNA ligase I have identified a nuclear localisation signal (NLS) within its N-terminal domain (Montecucco *et al.*, 1995). Moreover, a $p21^{CIP1}$-like PCNA binding site has been discovered in the extreme N-terminus of several eukaryotic DNA ligase I proteins, the investigation of which will be the subject of Chapter 4.

In recent years several reports have been published on the presence of DNA ligase proteins in mitochondria (see Chapter 1, section 1.2.3). In vertebrates such as *Xenopus* and humans, the DNA ligase III protein has been shown to function in mitochondria (Pinz and Bogenhagen, 1998, Lakshmipathy and Campbell 1999 and 2001) in addition to its roles in nuclear repair pathways and meiotic recombination (Mackey *et al.*, 1996, Nash *et al.*, 1997).

While this work was in progress, the DNA ligase I homologue of *S. cerevisiae* Cdc9, previously shown to be involved in nuclear DNA replication and repair pathways (see Chapter 1), was found to be present in the mitochondria also (Willer *et al.*, 1999, Donahue *et al.*, 2001). Localisation of Cdc9p to the different cellular compartments is regulated through the use of alternative translational start codons (Willer *et al.*, 1999). By mutational analyses Willer *et al.* showed that translation from the upstream AUG codon produced two mitochondrially-associated forms: a precursor and its cleavage product, the mature mitochondrial protein. Translation from a downstream start codon produced the major nuclear form of Cdc9. Mitochondrial protein extracts containing Cdc9 were able to seal nicks in double stranded DNA *in vitro* whereas extracts from a strain with a conditional lethal allele of *CDC9* were incapable of doing so (Donahue *et al.*, 2001). A functionally important role of

mitochondrial Cdc9 protein *in vivo* is indicated by the observations that a *cdc9*$^{ts}$ strain showed a reduced mitochondrial DNA content and had diminished capacity to repair DNA double-strand breaks, which were introduced by the inducible expression of mitochondrially targeted *EcoR*I (Donahue *et al.*, 2001).

In order to investigate whether *cdc17*$^{+}$ might also encode a mitochondrial protein, possible sub-cellular targeting functions of the N-terminal domain of Cdc17 were examined. Fusion proteins between the N-terminal 175 amino acids of Cdc17 and GFP were created and tested for their localisation. In this chapter evidence is presented that the N-terminal domain of Cdc17 is sufficient to target GFP to the nucleus and mitochondria. Amino acid residues that are essential for localisation to either compartment were identified through mutational analyses. In addition, the ability of mutated Cdc17 proteins with impaired targeting to nucleus or mitochondria are examined for the ability to rescue Cdc17 deficient strains.

## 3.2 Localisation studies

### 3.2.1 The N-terminal domain of Cdc17 targets GFP to nucleus and mitochondria

In order to establish if the N-terminal domain of Cdc17 could target the protein to the mitochondria, localisation studies using GFP-fusion proteins were performed.

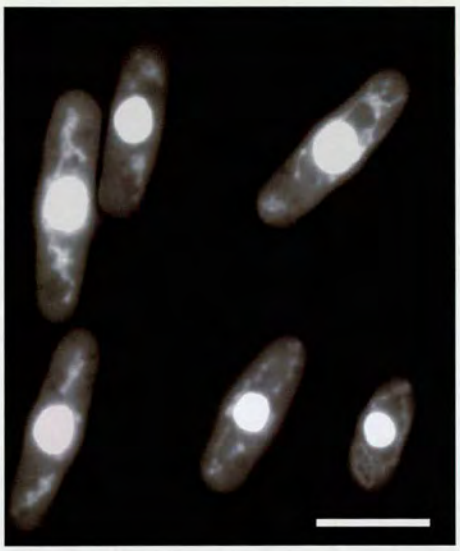
A portion of the *cdc17*$^{+}$ gene encoding the N-terminal domain and the first 50 amino acids of the CNCD was cloned into pREP3X-GFP in order to express a fusion protein consisting of GFP C-terminal to the first 175 amino acids of Cdc17 (see section 6.3.2.3). The resulting plasmid pREP3X-cdc17[1-175]-GFP was transformed into wild-type *S. pombe* cells (Sp5) and grown in EMM until mid-log phase in the absence of thiamine. Cells were harvested and examined by fluorescence microscopy (details see Materials and Methods, section 6.7.1 and section 3.2.5.1).

Figure 3.1 a) shows typical fluorescence patterns of cells expressing Cdc17[1-175]-GFP: bright staining throughout the whole nucleus was seen (sometimes the

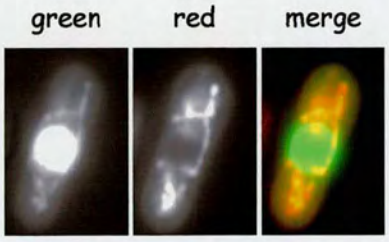
exclusion of fluorescence signals from the nucleolus is visible). Furthermore, other cellular structures exhibit fluorescence as well: extended tubules and string-like structures throughout the cytoplasm are visible. This staining pattern is similar to previously published images of *S. pombe* mitochondria (Robinow and Hyams, 1989, Yaffe *et al.*, 1996, Pelloquin *et al.*, 1999). To gain further evidence that the extra-nuclear fluorescence is due to localisation of the fusion protein to mitochondria, cells expressing the Cdc17-GFP-fusion were incubated with the mitochondrion-specific dye MitoTracker® Red (Molecular Probes) and the cells were viewed in the green and red channels for detection of GFP-fusion protein and MitoTracker respectively. Figure 3.1 b) shows double-labelling of a cell expressing the fusion with MitoTracker. Good colocalisation between extra-nuclear GFP fluorescence and the MitoTracker signal is apparent. In Figure 3.1 c) images of cells expressing GFP alone as a control are shown. Fluorescence is distributed throughout the whole cell including the nucleus.

In Summary, the N-terminal domain of Cdc17 (amino acids 1-175) is capable of targeting the heterologous GFP protein to both the nucleus and mitochondria.

a)



b)




c)

**Figure 3.1**: Images of *S. pombe* Sp5 cells expressing Cdc17[1-175]-GFP, a) and b), and GFP alone, c), from the induced *nmt1* promoter. Scale bar in panel a) represents 10 μm. In panel b) a cell expressing the GFP-fusion (left) is stained with the mitochondrion-specific dye MitoTracker® Red (middle). In the merged image on the right localisation of the GFP-fusion to mitochondria is indicated in yellow. Details of cell preparation and fluorescence microscopy are given in Materials and Methods, section 6.7.1.

### 3.2.2 Nuclear targeting

In order to investigate which residues within the N-terminal domain of Cdc17 are responsible for nuclear targeting, GFP-fusions of N-terminally-truncated versions of Cdc17[1-175] were examined for their localisation. pREP3X-GFP-vectors encoding the N-terminal domain of Cdc17 lacking the first 31, 65, 96 and 125 amino acids (see Materials and Methods) were expressed in *S. pombe* Sp5 cells and the distribution of GFP fluorescence determined. Cdc17[32-175]-GFP was capable of localising specifically to the nucleus whereas truncation of a further 34 amino acids (Cdc17[66-175]-GFP) resulted in non-specific fluorescence throughout the cell. Images of cells displaying typical fluorescence patterns for the different fusion proteins are shown in Figure 3.2 b) and c). It should be noted that the slightly increased fluorescence visible in the nucleus of cells expressing Cdc17[66-175]-GFP is unlikely to be specific since cells expressing only GFP exhibit a very similar fluorescence pattern (compare to Figure 3.1 c)). Cdc17[97-175]-GFP and Cdc17[125-175]-GFP also localised throughout the whole cell (data not shown). In conclusion, loss of amino acids 32-65 prevents the N-terminal portion of Cdc17 targeting GFP specifically to the nucleus.

On the other hand, amino acids 32-65 fused to GFP were deficient in specific nuclear localisation and displayed the same distribution as GFP alone (data not shown). Therefore, amino acids 32-65 are necessary but not sufficient for nuclear localisation. Examination of the amino acid sequence of the N-terminus of Cdc17 does not reveal any strong candidates for a classical nuclear localisation signal (NLS) (Nigg, 1997). However, residues 62-64 constitute a short stretch of basic amino acids, which is an integral part of an NLS. To test the role of these in nuclear targeting, a mutant was generated which was composed of the full length N-terminal domain where amino acids 62KRK64 were changed to 62AAA64. This GFP-fusion, Cdc17[1-175]ΔNLS-GFP, was expressed in *S .pombe* Sp5 from pREP3X. Figure 3.2 d) represents characteristic distribution of the fusion-protein. Cdc17[1-175]ΔNLS is unable to direct GFP specifically to the nucleus; green fluorescence is dispersed throughout the cell. The latter finding demonstrates that amino acids 62-64 of the N-terminus of Cdc17 are essential for specifically targeting the protein to the nucleus.

a) -cdc17[1-175]-GFP

b) -cdc17[31-175]-GFP

c) -cdc17[66-175]-GFP

d) -cdc17[1-175]ΔNLS-GFP

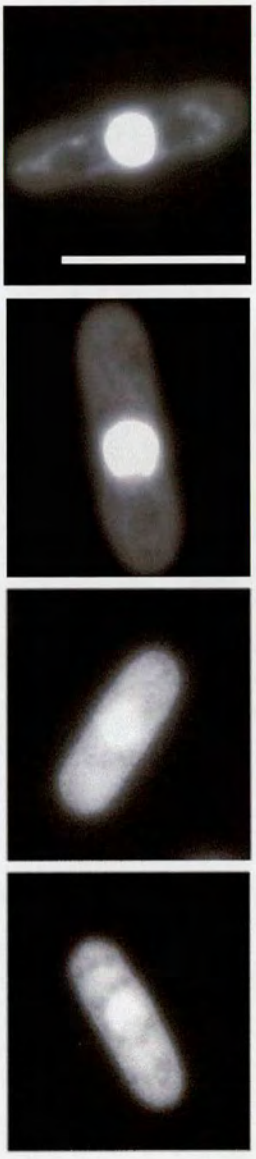


**Figure 3.2**: Images of *S. pombe* Sp5 cells transformed with pREP3X-vectors carrying the indicated inserts. Scale bar represents 10 μm. Cells were grown and prepared for fluorescence microscopy as described in Materials and Methods, section 6.7.1.

### 3.2.3 Mitochondrial targeting sequence of Cdc17

Amino acids 1-175 of Cdc17 are capable of targeting GFP to the nucleus and mitochondria (see section 3.2.1). Dissection of the N-terminal portion has elucidated amino acid residues that are essential for nuclear localisation (see previous section). What residues are essential for targeting to mitochondria? The fluorescence pattern of Cdc17[32-175]-GFP (see previous section, Figure 3.2 b)), which is exclusively nuclear, indicates that residues 1-31 are important for mitochondrial localisation.

The usage of alternative in-frame start codons for translation initiation to produce proteins which are each destined for a different cellular compartment is a common mechanism to generate multi-compartmentalised proteins from a single gene (reviewed in Danpure, 1995). Translation from the first methionine typically produces a mitochondrial precursor which is targeted to the mitochondria where it is processed to a mature mitochondrial protein. The amino acids encoded between the first and a second AUG codon often take the characteristic conformation of an amphiphilic α-helix, where one side of the helix predominantly contains positively charged residues while the opposite side is hydrophobic and uncharged (von Heijne, 1986). This characteristic presequence aids targeting to mitochondria, is recognised by the mitochondrial protein import machinery and is cleaved upon translocation through the mitochondrial membrane (Pfanner, 2000).

The extreme N-terminus of Cdc17 contains two in-frame methionine start codons at amino acid positions 1 and 20 (Figure 3.3 a)). Examination of the sequence between the two methionines revealed its potential to encode an amphipathic helix typical for mitochondrial targeting sequences (MTS). Figure 3.3 b) gives a schematic representation of the structure of the first 19 residues of Cdc17.

Scanning of the Cdc17 sequence using computer programmes aimed at identifying N-terminal sorting signals (Claros *et al.*, 1997) underlined the potential of Cdc17 to be imported into mitochondria. The PSORT II server (http://psort.nibb.ac.jp/) predicts the potential of a protein to be mitochondrial according to the method of Nakai and Kanehisa, 1992, which looks at the amino acid composition of the first 20 residues. It predicts that Cdc17 has properties which could enable it to be imported into mitochondria. The Mitoprot II 1.0 programme, which specifically examines the

potential of proteins to be targeted to mitochondria and/or chloroplasts (e-mail server: mitoprot@wotan.ens.fr), also predicts Cdc17 to be mitochondrial. In addition to the amino acid composition, this discriminant analysis takes into account several other parameters including the calculation of the maximal hydrophobicity of 7 neighbouring residues (this is the size of an α-helix thought to interact with membranes) on the maximal α-segment using four different scales of hydrophobicity (Claros and Vincens, 1996). The output of the calculations is the probability ($0<P<1$) of the protein to be imported into mitochondria given as DFM (discriminant function for mitochondrial proteins). Cdc17 has a value of 0.54, which is only slightly above the threshold (0.5) when proteins are considered to have properties to be imported into mitochondria. The other *S. pombe* DNA ligase proteins Lig4 and Lpp1 are unlikely to be targeted to mitochondria with DFM values of 0.06 and 0.05, respectively. Interestingly, the *Arabidopsis thaliana* DNA ligase I protein is found highly probable to be imported into mitochondria and/or chloroplasts with a value of 0.99. The *S. pombe* DNA replication proteins Pol3 (DFM=0.06), Rad2 (DFM=0.05) and Dna2 (DFM=0.07) are unlikely to be targeted to mitochondria. In contrast, *S. pombe* proteins known to be imported into mitochondria showed high prediction values, e.g. the Holliday junction resolvase Ydc2 (Doe *et al.*, 2000) with a probability of 0.81 and the UV-damage endonuclease Uve1 (Yasuhira and Yasui, 2000) with a predicted probability of 0.98.

a) .

1 $M_1$RTVFSQIPRFKQVNQYIR$M_2$STRQSDISNF 30

b)

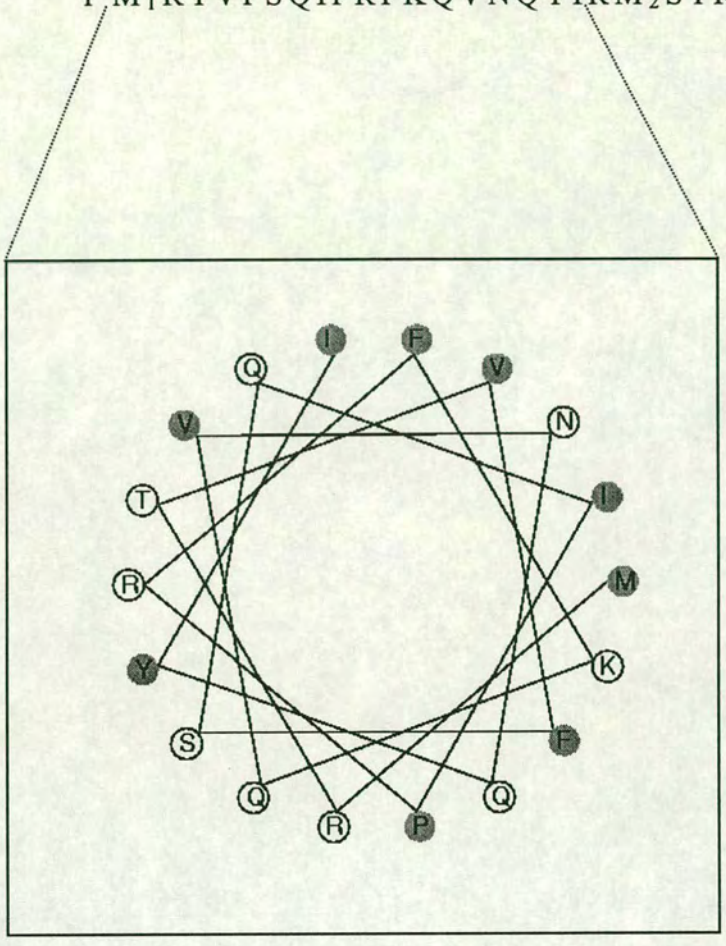



**Figure 3.3**: Putative mitochondrial targeting sequence of Cdc17. a) amino acid sequence of the first 30 amino acids of Cdc17. The first and second methionine is indexed 1 and 2, respectively. b) α-helical projection (http://marqusee9.berkeley.edu/kael/helical.htm) of the first 18 residues of Cdc17. Hydrophobic residues are shaded.

### 3.2.4 Targeting to mitochondria

The N-terminus of Cdc17 clearly has properties that allow targeting to mitochondria. If there were to be two forms of Cdc17 produced, one starting at the first methionine and harbouring the MTS (mitochondrial targeting sequence) and another one beginning at the second methionine at position 20 and therefore lacking the MTS, it can be reasoned that the latter form should constitute the nuclear version of Cdc17. In order to test this hypothesis, mutant forms of the N-terminal domain of Cdc17 fused to GFP were engineered and examined for their sub-cellular localisation.

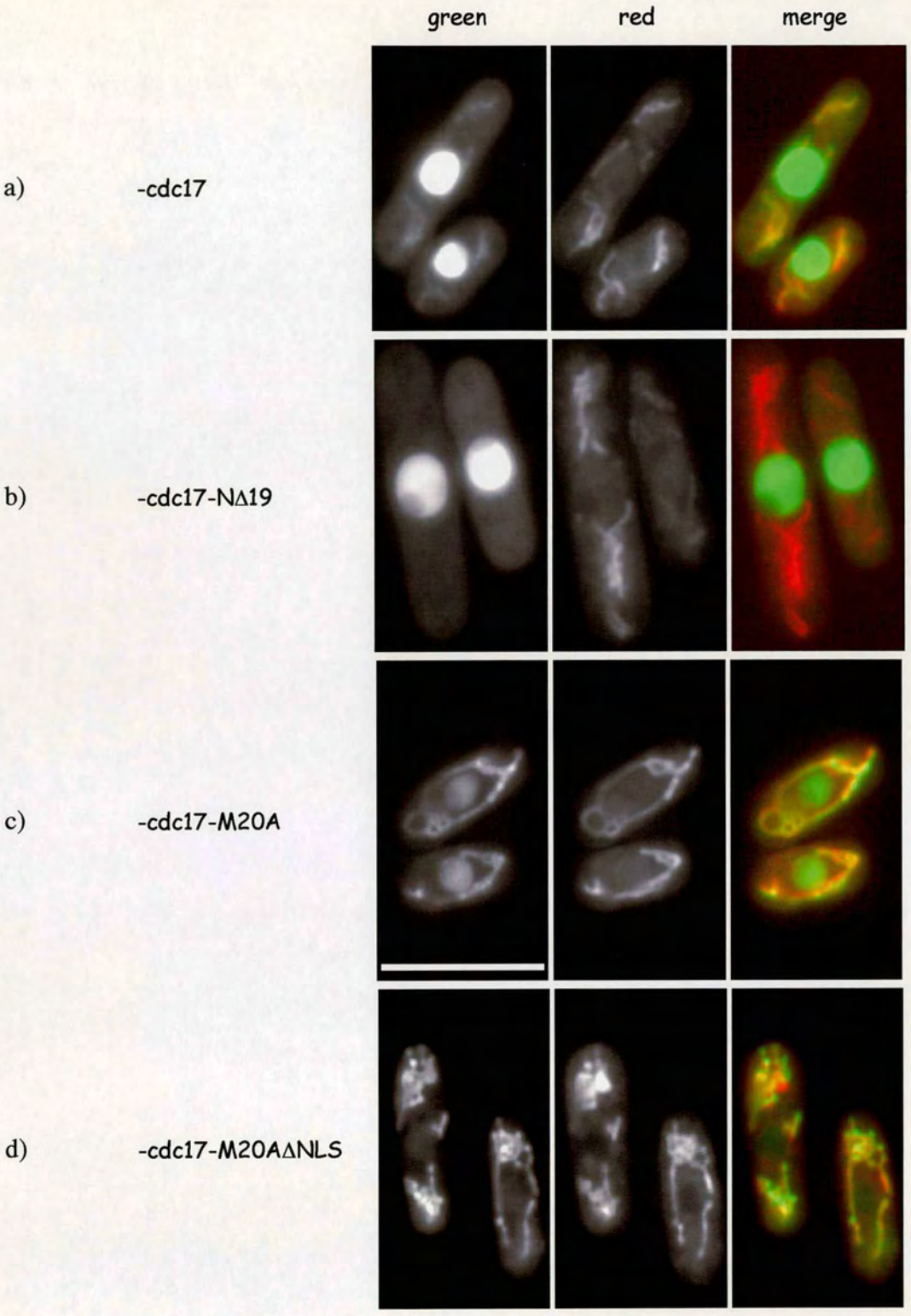
The GFP-fusion of the N-terminus of Cdc17 starting at the second methionine and therefore lacking the putative MTS, Cdc17[20-175]-GFP, was produced from the induced *nmt1* promoter in *S. pombe* Sp5 cells. Microscopic examination under the fluorescence microscope revealed solely nuclear localisation (Figure 3.4 b), left panel). The thread-like structures which represent mitochondria and are stained with MitoTracker Red (Figure 3.4 b), middle panel) are clearly absent in the green and merged image (the absence of yellow colour indicates that the fusion protein does not localise to mitochondria).

A second construct, consisting of the complete N-terminal domain and the beginning of the CNCD (amino acids 1-175) but where the second AUG – believed to be the translation initiation site for nuclear Cdc17 – is replaced with a GCG codon (Cdc17[1-175]M20A-GFP), was found to localise predominantly to mitochondria. The majority of cells displayed the staining pattern as seen in Figure 3.4 c). The fusion protein is still targeted to mitochondria as double labelling with MitoTracker Red clearly shows (Figure 3.4 c), right panel). However, nuclear staining is greatly diminished in most cases. Comparison of fluorescence intensity of cells expressing Cdc17[1-175]M20A-GFP, where the second methionine residue of the N-terminal domain of Cdc17 at position 20 is replaced by an alanine residue, to those expressing GFP-fusions of the wild-type, Cdc17[1-175], or the exclusively nuclear-targeted N-terminal domain, Cdc17[20-175], shows a drastic reduction in nuclear localisation: approximately 20% did not show any nuclear fluorescence and 80% displayed weak nuclear signals, which in intensity at most equalled mitochondrial fluorescence,

whereas in the former cases mitochondrial fluorescence sometimes was barely visible due to very bright nuclear fluorescence.

It was reasoned that the presence of the NLS, at least partly consisting of amino acids 62 KRK 64 (see section 3.2.2), in cells expressing Cdc17[1-175]M20A-GFP possibly allowed a small proportion of GFP-fusion-molecules to enter the nucleus despite the lack of the second start codon. To test this hypothesis, a third construct was engineered, in which the residues which had previously been identified as being essential for nuclear targeting, 62KRK64, were replaced with alanines in addition to the mutation of the second methionine. The fluorescence distribution of the resulting construct, Cdc17[1-175]M20AΔNLS-GFP is presented in Figure 3.4 d). If nucleotides encoding an integral part of the NLS, amino acids 62 KRK 64, are altered in addition to the second translation initiation start codon, the N-terminal domain of Cdc17 targets GFP to mitochondria only.

**Figure 3.4** (following page): Images of Sp5 cells expressing pREP3X-plasmids carrying the indicated inserts and stained with MitoTracker® Red (for details see Materials and Methods, section 6.7.1). The left panels show the intracellular distribution of the Cdc17-GFP-fusion and the middle panels represent mitochondria. The right panels show a merged image with the GFP-fusion in green, mitochondria in red and colocalisation of the GFP-fusion to mitochondria in yellow. The scale bar in panel c) represents 10 μm.


green
red
merge
a)
-cdc17
b)
-cdc17-NΔ19
c)
-cdc17-M20A
d)
-cdc17-M20AΔNLS

### 3.2.5 Further analyses of sub-cellular localisation

In the previous sections it was shown that the full-length N-terminal domain of Cdc17 (amino acids 1-175) directs GFP to the nucleus and mitochondria. Mutational analyses revealed residues important for targeting to the nucleus and mitochondria. In Appendix 4 a summary of engineered GFP-fusion proteins and their respective sub-cellular locations is given. Several additional experiments regarding the sub-cellular targeting functions of Cdc17 were carried out. These are described below.

#### 3.2.5.1 Cdc17-GFP localisation at reduced expression levels

Most of the fusion proteins described in the previous sections were also expressed in *S. pombe* Sp5 cells under the control of the weaker *nmt* promoter in pREP41X (promoter ON). The fluorescence distribution was indistinguishable from that seen with the pREP3X-constructs, albeit much weaker. Therefore, pictures were taken from cells expressing Cdc17[NT]-GFP from pREP3X. When Sp5 cells transformed with pREP3X-GFP-fusion vectors were grown in the presence of thiamine, i.e. promoter repressed, no fluorescence signals were detectable with the equipment used.

As was described in Chapter 2, it was observed that overexpression of *cdc17* from pREP3X (promoter ON) resulted in a *cdc* phenotype. In order to avoid any toxic effects on the cells due to overproduction of protein, the cells were kept in the presence of thiamine (promoter OFF) and only exposed to inducing conditions for a minimum amount of time (12-16 h) before viewing.

#### 3.2.5.2 Cdc17-GFP and replication foci

In an experiment aimed at investigating whether the N-terminus of Cdc17 – fused to GFP – formed foci within the nucleus, which would be indicative of sites of ongoing DNA replication, as has been described for human DNA ligase I (Cardoso *et al.*, 1997, Montecucco *et al.*, 1998, Leonhardt *et al.*, 2000), a time-course experiment

was performed. Sp5 cells transformed with pREP3X-cdc17[1-175]-GFP, grown in minimal medium containing thiamine, were shifted to thiamine-free medium and viewed at hourly intervals. However, no foci formation was detected. Instead, the fluorescence pattern went from no fluorescence after 9 hours incubation without thiamine to faint staining throughout the whole nucleus after 10 hours to intense nuclear staining thereafter. It cannot be concluded whether additional parts of the Cdc17 protein are required for sub-nuclear targeting of GFP or if it is simply that the resolution of the microscope and/or camera used was not sensitive enough.

### 3.2.5.3 Immunofluorescence studies

Affirmation that the respective full-length Cdc17 proteins localise in a similar fashion to Cdc17[1-175]-GFP-fusion-constructs came from preliminary immunofluorescence experiments. Sp5 cells transformed with pREP41X-cdc17, -cdc17-NΔ31 or -cdc17-M20A, grown in the absence of thiamine, were formaldehyde-fixed and spheroplasts were incubated with affinity-purified anti-Cdc17-peptide antibody (used at a 1:100 dilution) and stained with the secondary antibodies Alexa 488 or Cy3. Cells expressing full-length Cdc17 gave rise to nuclear and cytoplasmic signals. The cytoplasmic fluorescence was absent in Sp5/pREP41X-cdc17-Δ31 cells and no nuclear staining was apparent in cells expressing Cdc17-M20A. Figure 3.5 shows pictures of cells expressing the above described constructs. No signals were detected when the cells were grown in the presence of thiamine (promoter OFF) (data not shown).

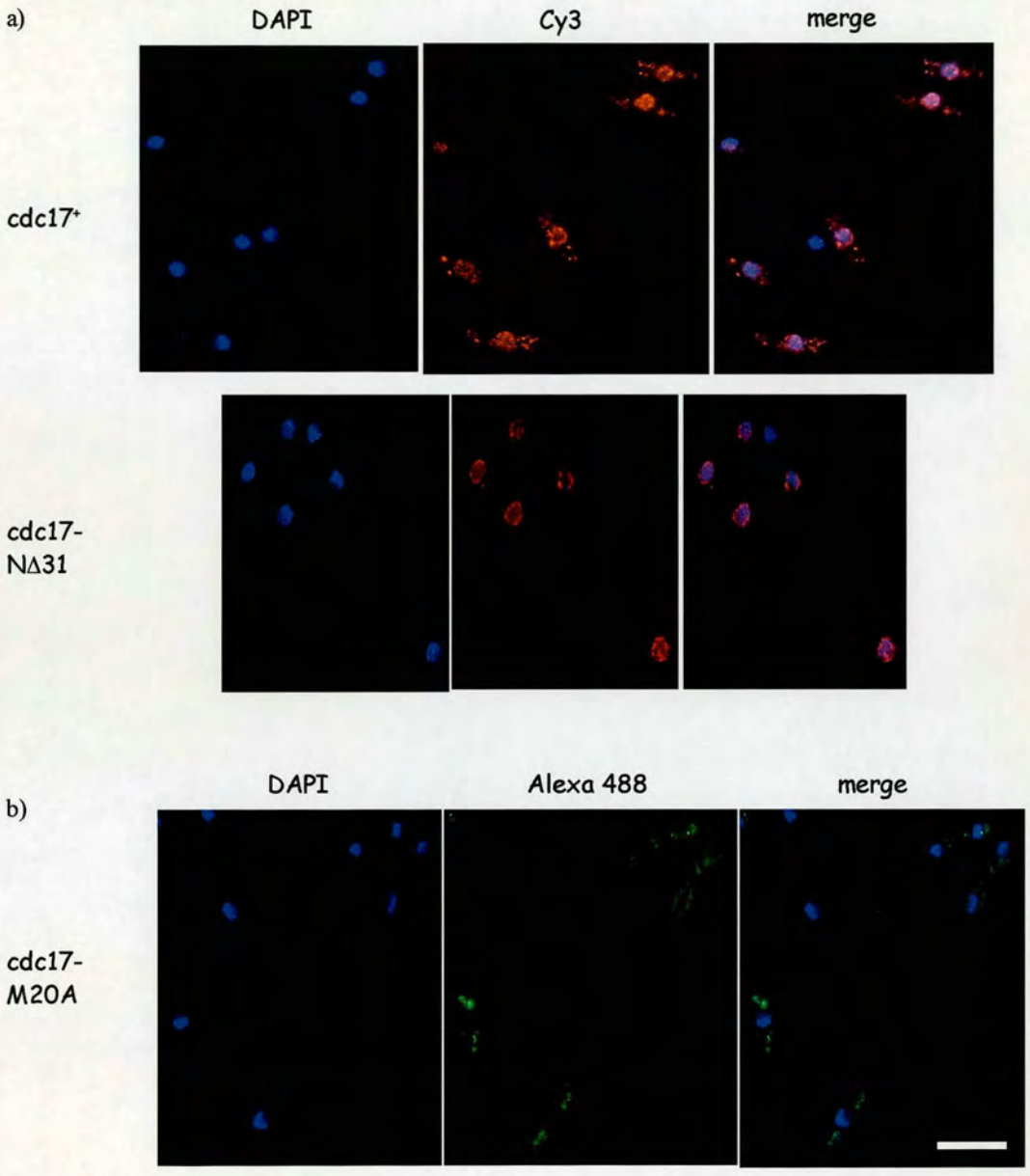



**Figure 3.5**: Immunofluorescence images of *S. pombe* Sp5 cells transformed with pREP41X-cdc17-constructs (as indicated on the left), grown in the absence of thiamine. Formaldehyde-fixed cells were incubated with affinity-purified @-Cdc17-peptide antibody (1:100 dilution) and subsequently in a) Cy3 (1:2000 dilution) or b) Alexa 488 (1:500 dilution) secondary antibody. Scale bar in b), right panel represents 10 μm. Details are described in Materials and Methods.

## 3.3 Detection of presumptive nuclear and mitochondrial Cdc17 proteins

Manipulations of the extreme N-terminus of Cdc17 result in altered targeting of GFP to the nucleus and/or mitochondria (see previous sections). In order to examine whether the presumptive nuclear (amino acid 20-768) and mitochondrial (amino acid 1-768 for the precursor form) forms of Cdc17 could be visualised by Western blotting, wild-type Sp5 cells harbouring pREP41X-cdc17, pREP41X-cdc17-NΔ19, pREP41X-cdc17-M20A and empty pREP41X vector were grown up to mid-logarithmic phase in the absence of thiamine and protein extracts were prepared from an equivalent amount of cells. The extracts were separated on an SDS-polyacrylamide gel and the Cdc17 proteins were Western blotted with affinity-purified anti-Cdc17-peptide antibody. The results are presented in Figure 3.6.

The full-length protein as well as the presumed nuclear Cdc17 protein starting from the second methionine at position 20, Cdc17-NΔ19, were produced to very high levels, i.e. 100-200 x the amount of endogenous Cdc17 protein (see section 2.4.4). In contrast, *cdc17* encoding the full-length protein with an alanine replacing the second methionine at position 20, Cdc17-M20A, was expressed at very low levels (Figure 3.6, compare lanes 1 and 3 to lane 2). This is consistent with the notion that the predominant (nuclear) form of Cdc17 is translated from the second AUG codon. Detection of Cdc17-M20A gave rise to two distinct bands (Figure 3.6 b), lane 2) and it is tempting to speculate that they represent the mitochondrial precursor and the mature mitochondrial form. The Cdc17- and Cdc17-NΔ19-proteins yield a number of different bands of similar molecular weight, the major one of which is of intermediate size to the protein bands visible in lane 2 and may well represent the nuclear form of Cdc17. Protein bands of slightly higher and lower molecular weight are seen with the full-length Cdc17 protein which are absent in Cdc17-NΔ19. These may correspond to the mitochondrial precursor and the mitochondrial form of Cdc17. Endogenous Cdc17 protein was not detectable under these conditions (see Figure 3.6, lane 4), which means that the visible protein species are solely the result of expression from pREP41X.

Similar results were obtained when protein extracts, prepared from Sp5 cells expressing the corresponding GFP-fusion proteins, were immunoblotted with anti-GFP antibody (data not shown). This confirms that the different expression profiles are most likely the result of alterations in the N-terminal domain of Cdc17.

Although an unambiguous assignment of protein bands to distinct nuclear and mitochondrial species is not possible, the immunoblotting analyses in Figure 3.6 (and data not shown), together with the localisation studies (section 3.2), are consistent with the conclusions that nuclear Cdc17 is translated from the second methionine and that the mitochondrial precursor form of Cdc17 starts from the upstream first methionine and subsequently is processed to the mature mitochondrial protein (see also sections 3.5 and 5.5).

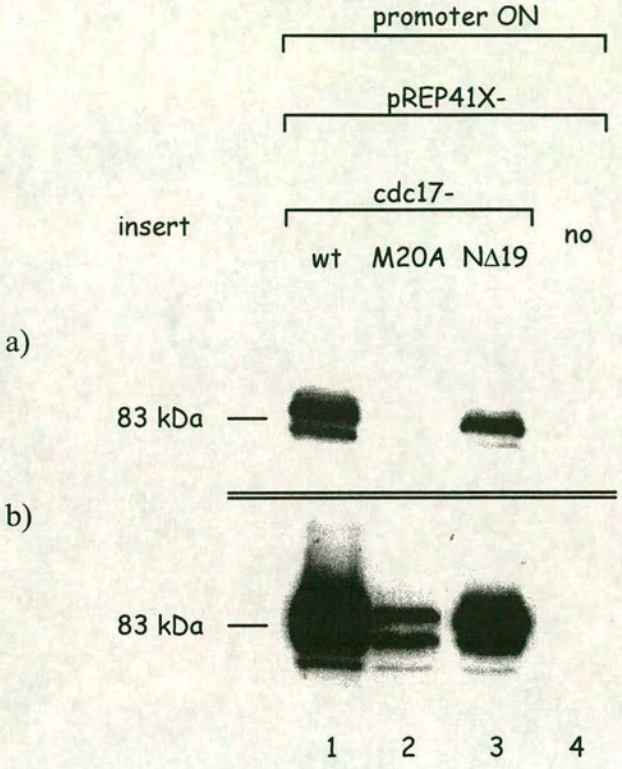



**Figure 3.6**: Western blotting analysis of protein extracts prepared from Sp5 cells transformed with pREP41X vectors containing the indicated *cdc17* constructs. The cells were grown in the absence of thiamine to mid-log phase and Western blotting was carried out as described in Materials and Methods. Cdc17 proteins were detected with anti-Cdc17-peptide antibody (1:200 dilution). a) short exposure, b) long exposure.

## 3.4 Rescue of *cdc17* deficient strains with N-terminally truncated Cdc17 proteins

### 3.4.1 Introduction

In the first chapter evidence was presented that Cdc17 lacking the N-terminal domain, if expressed to levels resembling wild-type, was unable to support growth of either the temperature-sensitive strain *cdc17-K42* or a *cdc17* null strain (section 2.5). Taking into account the results presented in the previous sections of this chapter, a likely explanation for this inability to complement *cdc17* deficient strains is the loss of intracellular sorting signals which lie within the N-terminal domain. In order to elucidate the functional relevance of the identified targeting signals, plasmids coding for Cdc17, Cdc17-NΔ19 (nuclear) and Cdc17-M20A (predominantly mitochondrial) were tested for functional complementation of *cdc17-K42*, *cdc17-L16* and *cdc17*Δ. In addition, the deletion strain was tested for rescue by additional N-terminally nested deletions.

### 3.4.2 Rescue of *cdc17$^{ts}$* strains

The functional complementation experiments of *cdc17$^{ts}$* strains were carried out as described in section 2.5.3 and Materials and Methods. Briefly, the *ts* strains transformed with the appropriate pREP41X plasmids were grown at 25°C and 1000 cells were spread on thiamine-containing selective media. The plates were incubated at 25°C and 36.5°C for 5 days. Incubation at 25°C in all cases resulted in excellent viability confirming that expression of the proteins did not exert any dominant negative phenotype. Figure 3.7 shows the results of the rescue experiments at 36.5°C. At the restrictive temperature, rescue of the temperature-sensitive strains by the predominantly mitochondrial Cdc17-M20A protein was quite similar for both temperature-sensitive alleles. Both strains were rescued reasonably well: the number of colony forming units was reduced by roughly 25% compared to rescue by the wild-type protein. The colonies formed varied considerably in size, a phenomenon

commonly seen when the rescued ts strains are incubated at the restrictive temperature and which likely reflects a variation in expression levels. Microscopic examination revealed that cells were elongated but while the cells were only moderately elongated in *cdc17-L16*, *cdc17-K42* cells showed a highly elongated phenotype.

The situation was entirely different for functional complementation by the Cdc17-NΔ19 protein. *cdc17-L16* was rescued to a level similar to wild-type. On the other hand, Cdc17-NΔ19 was unable to complement *cdc17-K42* at all. Microscopic examination showed the presence of microcolonies of round and swollen cells similar to those seen in *cdc17-K42*/pREP41X-NΔ125 (see Chapter 2, section 2.5.3). The experiments were also carried out where the proteins were moderately overproduced, i.e. induction of *nmt41* promoter through the absence of thiamine. In all but one case, rescue comparable to wild-type level was noted. Elevated levels of Cdc17-NΔ19, however, still failed to complement *cdc17-K42*.

In summary, a drastic difference is observed in the ability of the putative nuclear Cdc17-NΔ19 protein to functionally complement *cdc17-K42* and *cdc17-L16*.

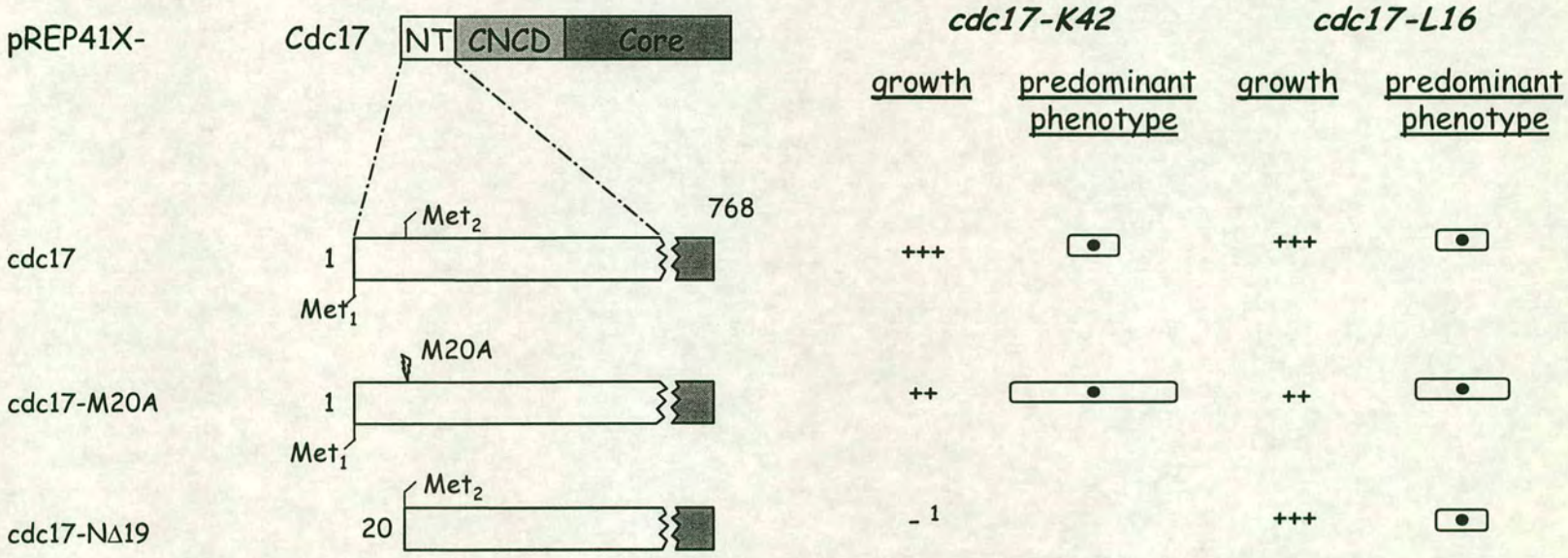



**Figure 3.7**: Ability of Cdc17 proteins to rescue the temperature-sensitive strains *cdc17-K42* and *cdc17-L16*. The ts strains, transformed with pREP41X-vectors carrying the indicated cdc17 constructs, were grown at 25°C and 1000 cells were plated onto EMM plates containing thiamine. Plates were incubated at 25°C and 36°C for 5 days. Rescue was scored from +++ (equivalent to wild-type) to – (no rescue). Predominant phenotype is indicated as wild-type-like (small cell), slightly elongated (medium cell) and *cdc* phenotype (large cell). 1: microcolonies of round and swollen cells.

### 3.4.2 Rescue of *cdc17*Δ

The functional complementation experiments of *cdc17*Δ using the plasmid shuffle approach were carried out as described in section 2.5.4 and Materials and Methods. Briefly, haploid *cdc17*Δ cells, kept alive by pUR19-cdc17 ($ura^+$), were transformed with pREP41X ($leu^+$) vectors harbouring different *cdc17* constructs (see Figure 3.8). 1000 cells containing both plasmids were plated out on EMM lacking leucine and containing thiamine and 5-FOA, and incubated at 32°C. Cells in arising colonies should only hold Cdc17 proteins expressed from pREP41X. The results are summarised in Figure 3.8.

After incubation for 4 days, only *cdc17*Δ cells expressing wild-type Cdc17 were able to grow. Neither expression of the predominantly mitochondrial Cdc17-M20A protein nor any of the N-terminal truncations yielded viable colonies (see Figure 3.8). Incubation for 7 days resulted in poor growth of deletion cells expressing a Cdc17 truncation lacking the N-terminal 65 amino acids (Cdc17-NΔ65), whereas no colonies were seen on any of the other plates (data not shown). If the plates were incubated for a further 3 days (total incubation time of 10 days) tiny colonies (<0.5 mm in diameter) appeared on the plates. Microscopic examination of these colonies revealed that, in most of the cases, cells were elongated to varying degrees compared to *cdc17*Δ rescued by the wild-type protein. The notable exception were *cdc17*Δ cells producing Cdc17-NΔ19, the protein starting from the second methionine – which is believed to constitute the nuclear form of Cdc17. In this case no cell elongation was visible (see Figure 3.8, right panel); this will be investigated further in the following chapter. The abortive microcolonies predominantly consisted of a small number of round and swollen cells, similar to those observed with *cdc17-K42* expressing Cdc17-NΔ125 (see section 2.5.3). Overproduction of all the N-terminally altered *cdc17* constructs to approximately 100-200 x the endogenous level functionally complemented *cdc17*Δ (data not shown).

The results presented show that loss of the first 19 amino acids of Cdc17 prevents the full rescue of a *cdc17*Δ strain. The data support the conclusion that mitochondrial DNA ligase function is essential for proliferation.

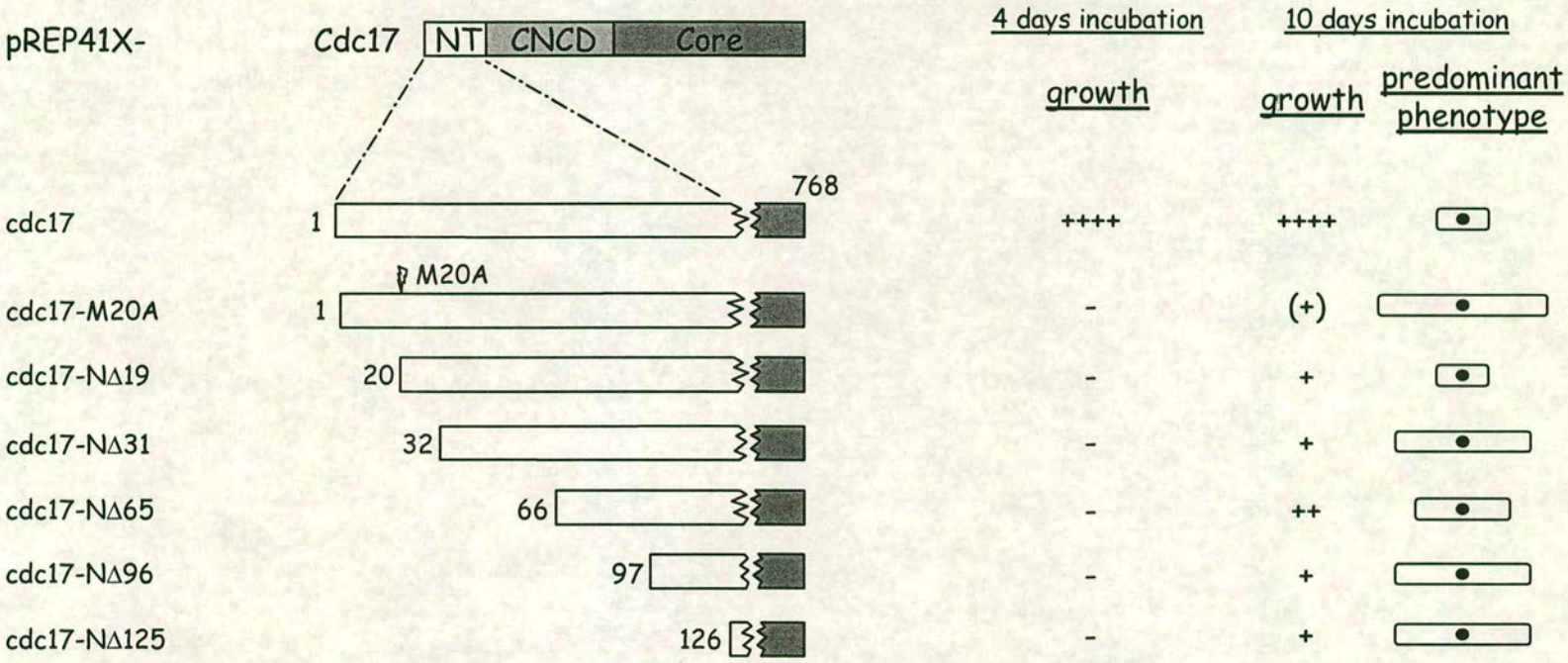



Figure 3.8: Rescue of *cdc17*Δ by N-terminally mutated Cdc17 proteins. The experiment was carried out as described in section 2.5.4 and Materials and Methods, section 6.6.11. Cdc17 proteins were expressed in the presence of thiamine (promoter OFF). Rescue was scored from ++++ (equivalent to wild-type) to – (no rescue). Predominant phenotype is indicated as wild-type-like (small cell), slightly elongated, elongated or highly elongated (large cell).

## 3.5 Discussion

This chapter provides evidence that the N-terminal 175 amino acids of the *S. pombe* DNA ligase I protein Cdc17 directs the heterologous GFP protein to the nucleus and mitochondria (section 3.2.1). Moreover, targeting signals to the respective organelles are contained within the first 65 amino acids of Cdc17 (section 3.2.2 and 3.2.3). Alternative usage of the in-frame methionine codons of Cdc17 in the extreme 5' region seems to be responsible for the dual localisation of the enzyme since lack of the sequence preceding the second methionine prevents mitochondrial targeting and modification of the second methionine drastically reduces nuclear localisation. Possible mechanisms whereby targeting to different cellular compartments may be achieved is discussed in section 5.5.

Preliminary immunofluorescence experiments have been carried out in order to ascertain whether the full-length proteins localise in a similar fashion to the N-terminal domain of Cdc17 (Figure 3.5). Generally, the localisation of full-length Cdc17 proteins with alterations in their N-termini correspond to that of the Cdc17[1-175]-GFP-fusions, i.e. expression of Cdc17 starting from the second methionine results in greatly reduced fluorescence of extra-nuclear structures and Cdc17 with a mutation in the second methionine does not yield nuclear fluorescence. It is, however, striking that nuclear fluorescence apparently concentrates on the nuclear periphery. It is not clear whether this represents true localisation to the nuclear rim or is an artefact, for example due to the primary antibody insufficiently penetrating the nucleus. To date, no reports on studies with yeasts are known, which describe the localisation of DNA replication foci or proteins to the nuclear periphery. For example, a recently developed technique to label sites of ongoing DNA replication by BrdU in yeast nuclei demonstrated that replication foci were distributed throughout the nucleus (Lengronne *et al.*, 2001). Also, in *S. pombe*, MCM subunits were found dispersed in the nucleus (Pasion and Forsburg, 1999). However, since, in the presented immunofluorescence experiments, expression of Cdc17 proteins with different N-termini produce different localisation patterns, the results are considered significant.

Investigation of residues required for nuclear targeting revealed that amino acids 62 KRK 64 are necessary for targeting function. They are however, not sufficient since amino acids 32-65 alone are incapable of specific nuclear sorting. Previous studies on conserved features of nuclear localisation signals have identified two classes: the first one is the basic-type NLS composed of a few basic residues that are either monopartite of the consensus $(K/R)_{4-6}$ or bipartite with the consensus motif $(K/R)_2$ $X_{10-12}$ $(K/R)_3$ (Nigg, 1997). Other types of NLS which show no obvious resemblance to the classical NLS and are heterogeneous in nature have been determined experimentally. In a screen for novel NLS, clusters of 2-3 arginine or lysine residues that were too close or too far apart to be classified as the bipartite basic NLS were identified (Christophe *et al.*, 2000). With respect to Cdc17, close inspection of the N-terminal sequence of Cdc17 revealed another small cluster of basic residues: 97 KKQK 100. Possibly, this serves as the second part of the NLS of *S. pombe* Cdc17.

In order to visualise differences in the presumed nuclear and mitochondrial forms of Cdc17, protein extracts of cells producing the distinct forms were detected with anti-Cdc17 antibody. Immunoblotting analyses of wild-type Sp5 cell extracts moderately overexpressing Cdc17 identified various Cdc17 species. Expression of the N-terminally truncated protein Cdc17-ΔN19 yielded product levels similar to that of the wild-type gene and also resulted in a smear of several indistinct bands. It cannot be concluded if the proteins detected are caused by physiological or artefactual degradation, possibly due to overproduction, or if they reflect post-translational modifications. Attempts were made to separate cell extracts into a mitochondrial fraction and a post-mitochondrial supernatant and detect the respective associated Cdc17 proteins. However, the lack of an antibody to an *S. pombe* mitochondrial protein suitable for immunoblotting, greatly hindered this task since the successful separation of sub-cellular fractions could not be monitored.

Complementation of the *cdc17*$^{ts}$ strains *cdc17-K42* and *cdc17-L16* with Cdc17-NΔ19, which is the presumed nuclear form of the enzyme, elucidated further differences in the nature of the temperature-sensitive Cdc17 proteins (already

discussed in section 2.6). The inability of Cdc17-NΔ19 to rescue *cdc17-K42* may reflect the possibility that thermolabile Cdc17-K42 is also defective in a function required for mitochondrial DNA ligase activity. Therefore, additional production of nuclear Cdc17 cannot compensate for the requirement of mitochondrial Cdc17 function. Whenever *cdc17-K42* or *cdc17*Δ was complemented with low levels of Cdc17 proteins deleted for the N-terminal 19 to 125 amino acids (which lack the mitochondrial targeting sequence), the majority of cells were unable to produce colonies and microscopic examination showed accumulation of 16-<500 round and swollen cells. The remainder display an extremely slow growth phenotype. These phenotypes are presumably a consequence of defective Cdc17 function in mitochondria.

## 3.6 Summary

This chapter presents evidence that the N-terminal 175 amino acids of Cdc17 target the heterologous GFP protein to the nucleus and mitochondria. The full-length Cdc17 protein also localises to both cellular compartments. The N-terminal 19 amino acids are essential for targeting GFP to mitochondria since their deletion abolishes localisation to this organelle but still allows targeting to the nucleus. The major nuclear form of Cdc17 is probably translated from the second in-frame AUG codon encoding for a methionine at position 20, mutation of which results in greatly reduced nuclear localisation. Furthermore, residues essential for sorting to the nucleus were also identified. Immunoblotting analyses showed that proteins localising to the nucleus produce a major species which is of intermediate size to two protein species, which are present at >10-fold lower levels and are being produced as the result of expression of a *cdc17* gene where the second AUG codon had been mutated. Cdc17 function is required in both organelles since a *cdc17* deletion strain cannot be fully complemented by expression of the presumed nuclear or mitochondrial Cdc17 proteins.

# 4. Role of PCNA binding motif for *in vivo* function of Cdc17

## 4.1 Introduction

The *S. pombe* DNA ligase I homologue Cdc17 contains a PCNA binding site at its N-terminus: amino acids 24 to 31 (24 QSDISNFF 31) constitute an evolutionarily conserved PCNA binding motif which has first been identified in the mammalian cell cycle and DNA replication inhibitor $p21^{CIP1}$ (Xiong *et al.*, 1993, Warbrick *et al.*, 1995) (see Introduction, section 1.1.5). Preliminary experiments have confirmed that the N-terminal 175 amino acids of Cdc17 are able to bind the *S. pombe* PCNA homologue Pcn1: Cdc17[1-175]-GST bound to recombinant His-tagged Pcn1 in pull-down experiments (Stuart MacNeill, personal communication). Most proteins containing such a motif, which have been tested for their ability to interact with PCNA *in vitro*, have been found to do so. However, several studies that were aimed at investigating the functional relevance of the interaction of PCNA with the target protein *in vivo* have given rise to varying results. The PCNA binding motif in the C-subunit Cdc27 of *S. pombe* DNA polymerase δ has been shown to be essential for function of the protein *in vivo* (Reynolds *et al.*, 2000). Indeed, a very recent report suggested that the p21-like PCNA binding motif present in the human homologue of pol δ C-subunit p66 constitutes the major PCNA binding site required for PCNA-dependent stimulation of Pol δ in DNA replication (Ducoux *et al.*, 2001). On the other hand, the canonical motif is also present in the N-terminus of large subunit Rfc1 of the five-subunit clamp loader RF-C. An N-terminal deletion including the motif of *S. cerevisiae* Rfc1 however apparently did not affect the replication function of RF-C (Gomes *et al.*, 2000) and a similar result was obtained with an N-terminal deletion in the human Rfc1 homologue in *in vitro* assays (Uhlmann *et al.*, 1997).

In the previous chapter evidence was presented that $cdc17^{+}$ encodes nuclear and mitochondrial forms of DNA ligase I. The extreme N-terminus appears to be responsible for targeting the enzyme to mitochondria. It was hypothesised that a mitochondrial targeting sequence at its N-terminus directs Cdc17 to the

mitochondrial import machinery where the presequence is cleaved off upon translocation through the mitochondrial membranes to yield the mature mitochondrial protein. Predictions of the cleavage site range from position 21 (using the MitoProt programme) to position 33 (PSORT). Mitochondrial Cdc17 appears to perform important functions as judged by the inability of a truncated Cdc17 protein, Cdc17-NΔ19 to support normal growth of a *cdc17*Δ deletion strain (see Chapter 3, section 3.4).

In this chapter an attempt is undertaken at investigating the relevance of the PCNA binding motif and therefore, by implication, the importance of binding to PCNA for *in vivo* function of nuclear DNA ligase I in *S. pombe*. Taken into consideration that the PCNA binding motif and the mitochondrial presequence cleavage site are potentially overlapping it becomes clear that a straightforward approach is inadequate for this task, as any manipulation of the motif may affect the mitochondrial targeting function. Evidence is presented that it is possible to separate the two functions of Cdc17 by expressing distinct nuclear and mitochondrial forms. This strategy is used to manipulate only the nuclear Cdc17 protein with the aim of elucidating the significance of the PCNA binding motif for *in vivo* function. Firstly, further physiological analyses are presented which point toward a functionally important role of the motif.

## 4.2 Genetic analyses

Inhibition of cell cycle progression, e.g. due to defects in DNA replication, is reflected by elongation of *S. pombe* cells. In Chapter 2 it was shown that overexpression of Cdc17 caused wild-type cells to elongate by about 60% (22.3 μm for cells overexpressing *cdc17*$^{+}$ compared to 13.9 μm for control cells, see section 2.4). In order to determine if the presence of the PCNA binding motif contributes to the cell cycle delay seen under these circumstances, a *cdc17* construct which is deleted for sequences encoding the N-terminal 31 amino acids including the p21$^{CIP1}$-like PCNA binding motif (24 QSDISNFF 31) was cloned 3' to the *nmt1* promoter and overexpressed in *S. pombe* Sp5 cells. In the absence of the PCNA binding motif the cell elongation was significantly reduced: cells divided at a cell length of (17.6 ± 4.7) μm which corresponds to an increase of only 27% relative to cells transformed with empty vector.

Additional indications that overproduction of proteins containing the PCNA binding motif causes deleterious effects on cell cycle progression came from three other observations. Firstly, Sp5 cells grossly overproducing Cdc17-NΔ19, which possesses the motif and is believed to constitute the nuclear form of Cdc17, divided at an increased cell length of (21.4 ± 6.3) μm. Secondly, cells expressing the N-terminal domain plus the conserved non-catalytic domain of Cdc17 at high levels (i.e. Sp5 cells transformed with pREP3X-cdc17-CΔ379, grown in the absence of thiamine) were highly elongated with a cell size at division of (21.1 ± 6.2) μm, which is comparable to the cell size seen with cells overproducing the full length protein ((22.3 ± 7.4) μm). Thirdly, in order to confirm that cell cycle delay is caused directly by high levels of a protein carrying the PCNA binding motif, which is likely to sequester the PCNA molecules in the cell, another construct was generated. Sequences encoding amino acids 1-31 of Cdc17, which include the mitochondrial targeting sequence and the motif, were replaced with sequences encoding the p21$^{CIP1}$-like PCNA binding motif of the *S. pombe* DNA replication protein Rfc1 (1 MSNSDIRSFF 10). This construct was cloned into pREP41X, resulting in pREP41X-cdc17-NΔ31-rfc1[1-10], and introduced into Sp5 cells. Transformants

were grown in the presence (promoter OFF) and absence (promoter ON) of thiamine and cell length was determined. Overproduction of this construct resulted in a cell elongation of approximately 40% compared to low-level-expression (20.7 μm versus 14.9 μm). Note that the moderate overproduction of Cdc17 with the PCNA binding motif of Rfc1 has a much more drastic effect on cell cycle progression than expression of the full-length wild-type Cdc17 protein. Western blotting confirmed that both constructs were expressed to comparable levels (data not shown). Table 4.1 summarises the cell length measurements and Figure 4.1 gives an overview of the results.

**Table 4.1**: Cell lengths of *S. pombe* Sp5 cells carrying the indicated plasmids. Cells were grown to mid-log phase either in the presence (promoter OFF) or absence (promoter ON) of thiamine and cell length was determined.

| plasmid | promoter OFF | promoter ON |
|---|---|---|
| pREP3X-cdc17-NΔ31 | --- | (17.6 ± 4.7) μm |
| pREP3X-cdc17-NΔ19 | --- | (21.4 ± 6.3) μm |
| pREP3X-cdc17-CΔ379 | --- | (21.1 ± 6.2) μm |
| compare: pREP3X-cdc17 | --- | (22.3 ± 7.4) μm |
| pREP41X-cdc17-NΔ31-rfc1[1-10] | (14.9 ± 1.3) μm | (20.7 ± 7.0) μm |
| compare: pREP41X-cdc17 | (15.1 ± 1.4) μm | (15.1 ± 3.3) μm |
| pREP41X | (15.0 ± 1.6) μm | (14.7 ± 1.8) μm |

In the previous chapter results were presented regarding the ability of N-terminal truncated Cdc17 proteins to rescue a *cdc17* deletion strain. Under normal conditions no rescue was observed with proteins truncated for the first 19 and 31 residues respectively, most likely due to defect of the mitochondrial function of Cdc17, which could be the cause of severely compromised growth. However, much prolonged incubation resulted in the appearance of tiny colonies. Microscopic examination of these revealed that *cdc17*Δ cells expressing a Cdc17 protein which lacks the first 19 residues, but still possesses the PCNA-binding motif, were of a size equivalent to wild-type whereas *cdc17*Δ cells expressing Cdc17 lacking the PCNA binding motif, Cdc17-NΔ31, were highly elongated. The above data point toward an important role

of the PCNA binding motif for *in vivo* function, suggesting that nuclear function of Cdc17 depends on an interaction with PCNA.

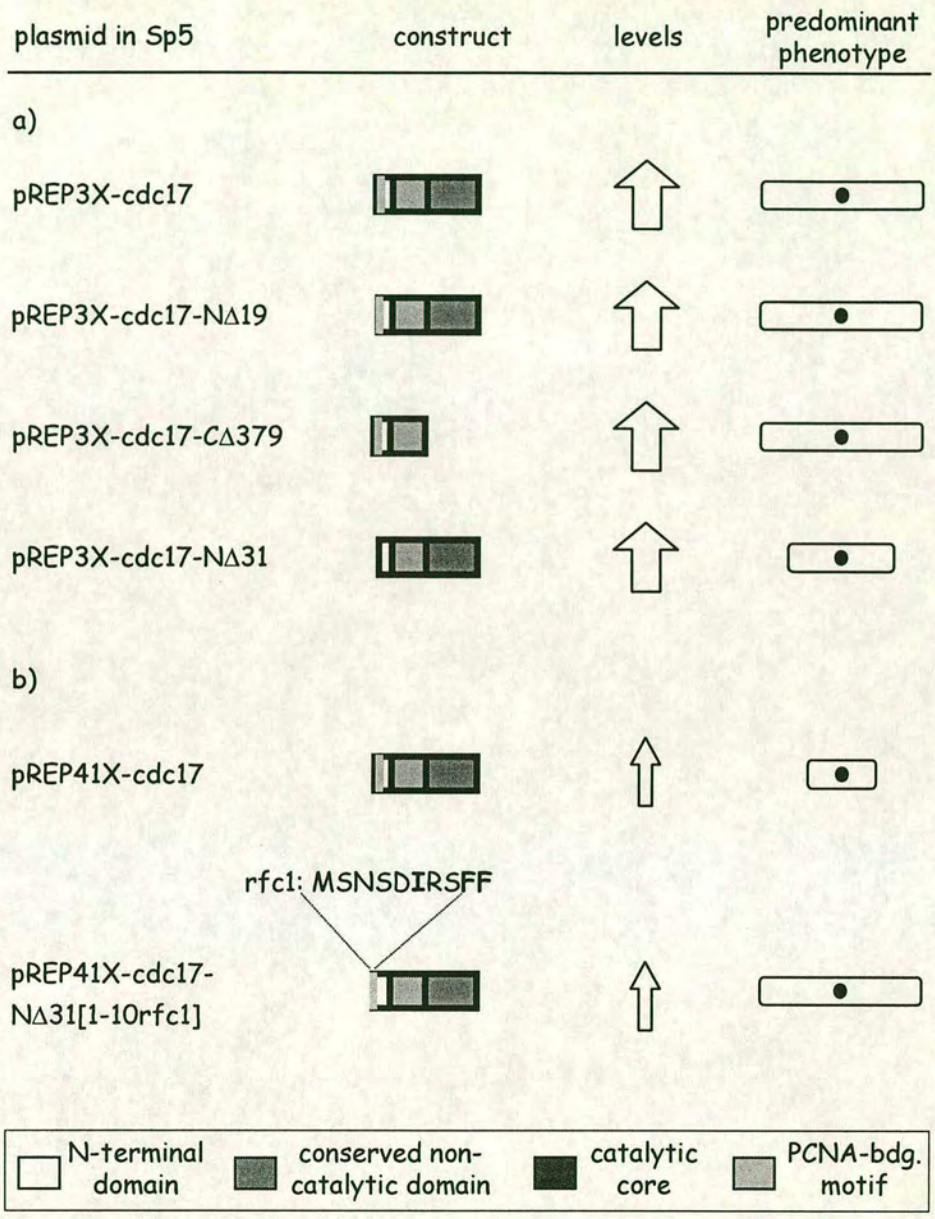



**Figure 4.1**: Summary of cell length measurements. *S. pombe* Sp5 cells, transformed with the indicated plasmids were grown to mid-log phase in the absence of thiamine and cell length at division was measured. Different degrees of cell elongation are shown. Thick arrows denote expression from the induced *nmt1* promoter and thin arrows indicate expression from the induced *nmt41* promoter. See text for details.

## 4.3 Separate expression of nuclear and mitochondrial Cdc17

### 4.3.1 Introduction

Evidence presented in the previous section indicates that the PCNA binding motif of Cdc17 may carry out essential functions. It would be of great interest to find out if the *in vivo* function of nuclear Cdc17 was dependent on an interaction with PCNA. Since the motif potentially overlaps with the mitochondrial presequence, its manipulation for the purpose of investigating its function is likely to impair mitochondrial Cdc17 function. Studies on the intracellular targeting of GFP by different versions of the N-terminal domain of Cdc17 revealed that some constructs localised exclusively to the nucleus whereas others were predominantly targeted to mitochondria (see Chapter 3, Figure 3.4). If it was possible to simultaneously produce Cdc17 proteins that are targeted to one cellular compartment only, i.e. the nucleus or mitochondria, then the dual requirement for functions of Cdc17 should be satisfied. This would then allow the manipulation of the PCNA binding motif in the nuclear protein without perturbing mitochondrial DNA ligase function. In order to determine if separate expression of genes encoding nuclear and mitochondrial Cdc17 proteins can successfully substitute for expression of the wild-type gene a strategy for co-expression of the respective proteins in *cdc17*Δ was devised.

### 4.3.2 Experimental strategy

A plasmid shuffle system was described in Chapter 2 which was used to investigate the ability of truncated or mutated Cdc17 proteins to rescue the deletion strain (section 2.5). Based on this approach a method for the simultaneous expression of the nuclear protein Cdc17-NΔ19 and the predominantly or exclusively mitochondrial protein Cdc17-M20A and Cdc17-M20AΔNLS, respectively, was designed. A prerequisite for this double plasmid shuffle approach is the observation that moderate overexpression of N-terminally altered Cdc17 proteins are able to sustain growth of a strain deleted for *cdc17* (see Chapter 3, section 3.4). In contrast, expression of either

Cdc17-NΔ19, Cdc17-M20A or Cdc17-M20AΔNLS at levels similar to wild-type, was unable to rescue the *cdc17* deletion strain (see previous Chapter, section 3.4 and Figure 3.8 and data not shown).

The experimental strategy of choice is outlined in Figure 4.2 a) and details are described in Materials and Methods. Briefly, the leu$^-$ ura$^-$ haploid strain *cdc17::kanMX6 leu1-32 ura4-D18 his7-366 ade6-M21?*, in which the only *cdc17* allele is replaced with the *kanMX6* marker, is maintained by a wild-type *cdc17* copy expressed from pUR19 (ura$^+$). The plasmid pREP41X-cdc17-NΔ19 encoding nuclear Cdc17 is introduced into that strain and plated out on media selecting for both plasmids. A well-isolated transformant was then grown up in liquid culture and plated onto minimal medium selecting for the pREP41X plasmid (leu$^+$) and containing the drug 5-FOA, which selects against ura$^+$ cells. At this stage the media lacks thiamine which allows the production of elevated levels of Cdc17-NΔ19. After confirmation that the strain is ura$^-$, i.e. that pUR19-cdc17 is lost, the pREP42 (ura$^+$) vector carrying *cdc17-M20A* or *cdc17-M20AΔNLS*, encoding for mitochondrial Cdc17, was transformed in. In a last step, *cdc17*Δ cells harbouring pREP41X-cdc17-NΔ19 and pREP42X-cdc17-M20A (or -M20AΔNLS) was grown up and 1000 cells were plated on selective media containing thiamine (i.e. promoter OFF).

After incubation for five days at 32°C >900 colonies were visible. Microscopic examination of the cells confirmed that the cell morphology was similar to wild-type cells with a cell length of (14.4 ± 3.3) μm (see Figures 4.2 b), right panel and Figure 4.3, left panel). As a positive control pREP41X-cdc17 was introduced instead of pREP41X-cdc17-NΔ19. The colony number was comparable, the only difference being a more uniform colony size range. Negative controls included empty pREP41X vector in the first plasmid shuffle (i.e. lack of nuclear Cdc17), which did not produce any viable colonies, and the introduction of empty pREP42X in the second shuffle ( i.e. lack of mitochondrial Cdc17), which also resulted in no colonies under these experimental conditions. The reason why Cdc17-M20A and Cdc17-M20AΔNLS were both used as mitochondrial proteins will be commented on at a later stage. At this point it is only worth noting that both can serve as a source of mitochondrial Cdc17.

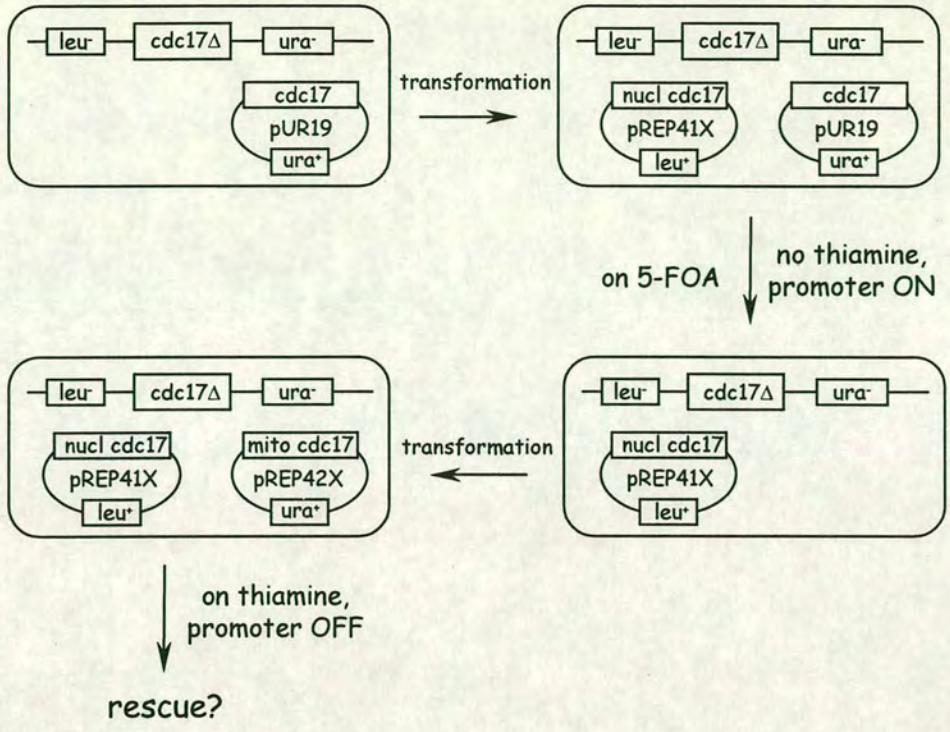
Figure 4.2 b) summarises the result together with previous results. The ability of the simultaneous expression of Cdc17-NΔ19 and Cdc17-M20A or Cdc17-M20AΔNLS to restore viability of *cdc17*Δ shows that nuclear and mitochondrial Cdc17 functions can be separated.

From now on the designations outlined in Table 4.2 will be used in order to facilitate comprehensibility.

**Table 4.2**: Designations used throughout the remainder of this chapter.

| designation | explanation | |
|---|---|---|
| *cdc17*Δ | *cdc17::kanMX6 leu1-32 ura4-D18 his7-366 ade6-M21?* | |
| low expression | expression from pREP41X in the presence of thiamine, yields approximately endogenous Cdc17 levels | |
| overexpression | expression from pREP41X in the absence of thiamine, 100 – 200-fold overproduction | |
| nuclear Cdc17 | Cdc17-NΔ19<br>possesses PCNA-bdg. motif<br>**or** Cdc17-NΔ31<br>lacks PCNA-bdg. motif | genes encoding these proteins are expressed from pREP41X ($leu^+$) |
| mitochondrial Cdc17 | Cdc17-M20A<br>(predominantly mitochondrial)<br>**or** Cdc17-M20AΔNLS<br>(exclusively mitochondrial) | genes encoding these proteins are expressed from pREP42X ($ura^+$) |

a) Double plasmid shuffle: co-expression of nuclear and mitochondrial Cdc17




b) Co-expression rescues *cdc17*Δ.

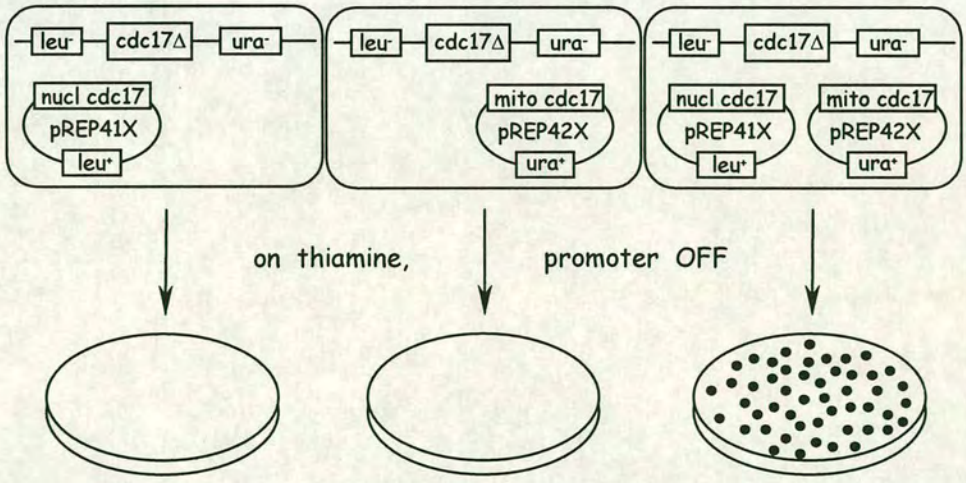



**Figure 4.2**: a) Schematic representation of the double plasmid shuffle approach used to separately express nuclear and mitochondrial Cdc17. b) Co-expression of nuclear and mitochondrial Cdc17 sustains viability of cdc17Δ cells. nucl = nuclear (encoding Cdc17-NΔ19), mito = mitochondrial (encoding Cdc17-M20A or Cdc17-M20AΔNLS)

### 4.3.3 Effect of removing the PCNA binding motif from nuclear Cdc17

In a next step, the strategy described above was used to assess the importance of the presence of the PCNA binding motif. A vector encoding nuclear Cdc17 lacking the PCNA binding motif was introduced into *cdc17*Δ/pUR19-cdc17 and the double plasmid shuffle was performed, as outlined in Figure 4.2 a) and described in the previous section. Expression of nuclear Cdc17 possessing the motif served as a positive control. Initially, Cdc17-M20A was used as the mitochondrially targeted protein. 1000 cells, expressing nuclear Cdc17 with or without the motif as well as Cdc17-M20A, grown in the absence of thiamine, were plated out on selective media containing thiamine and incubated for five days. The results are summarised in Figure 4.3. Loss of the PCNA binding motif in the nuclear Cdc17 protein caused a drastic decrease in cell viability as judged by the number of colony forming units (see photographs in Figure 4.3). Additionally, cells were highly elongated compared to the control (32.2 μm versus 14.4 μm). Note that the size of the existing colonies varied considerably.

The conclusion from this experiment is that the presence of the PCNA binding motif is important for nuclear DNA ligase I function. It was speculated that those 15% of cells that were able to form colonies could do so because a fraction of the mitochondrial protein Cdc17-M20A, which retains the PCNA binding motif as part of the putative mitochondrial presequence, was somehow capable of performing nuclear Cdc17 function. Consistent with this assumption would be the observation that – in the intracellular localisation studies – Cdc17[1-175]-M20A-GFP showed some nuclear fluorescence (see Chapter 3, Figure 3.4 c)). In order to investigate this possibility, it was decided to employ the Cdc17-M20AΔNLS protein as mitochondrial Cdc17. The corresponding N-terminal domain had targeted GFP exclusively to mitochondria (see Chapter 3, Figure 3.4 d)). Also it had been shown that Cdc17-M20AΔNLS was able to rescue *cdc17*Δ when co-expressed with Cdc17-NΔ19 (see previous section, legend to Figure 4.2).

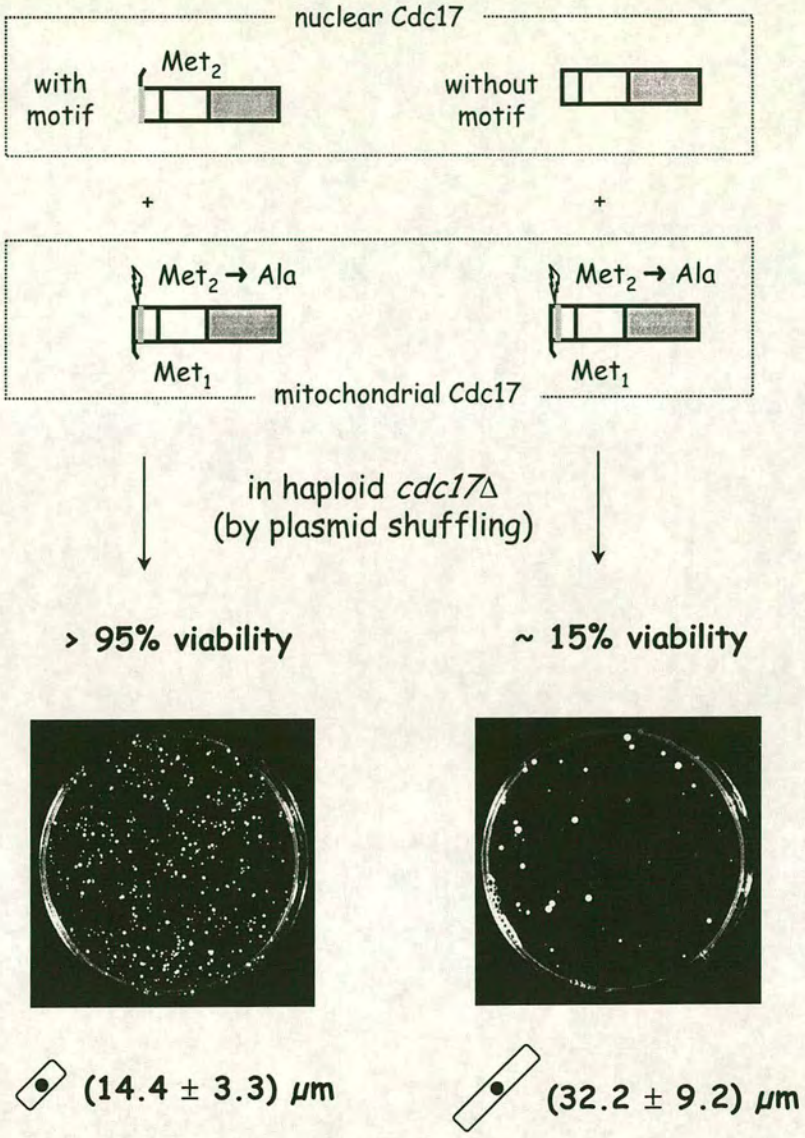



**Figure 4.3**: Result of the double plasmid shuffle experiment, explained in the text. Nuclear proteins are Cdc17-NΔ19 (with motif) and Cdc17-NΔ31 (without motif) and the depicted mitochondrial protein is Cdc17-M20A. *cdc17*Δ cells containing both plasmids were plated out on selective media containing thiamine and incubated at 32°C for five days.

Using the double plasmid shuffle approach the ability of nuclear Cdc17 lacking the PCNA binding motif and Cdc17-M20AΔNLS to sustain growth of *cdc17*Δ was investigated. However, no significant difference was detected relative to the use of Cdc17-M20A as the mitochondrial protein: the experiment was carried out several times and the number of colonies that arose lay between 10% and 30% compared to the control (Cdc17-NΔ19 as nuclear protein). Again, the rescued cells were considerably elongated compared to control cells (see Figure 4.4 top panels).

It is possible that the survival of a proportion of *cdc17*Δ cells expressing nuclear Cdc17 lacking the PCNA binding motif and a mitochondrial form may be due to high levels of these proteins being present in the cells (bearing in mind that overexpression of those proteins rescues the deletion strain). In order to investigate this possibility colonies from the selective plates containing thiamine were re-streaked on thiamine or grown up in liquid culture before plating out on thiamine a second time. However, it was found that cells were able to grow when re-streaked on thiamine. Similarly, cell material taken from medium- to large-diameter colonies did grow in liquid culture containing thiamine, although 4-5 times slower than *cdc17*Δ cells expressing nuclear Cdc17 possessing the motif alongside mitochondrial DNA ligase. When an equivalent number of cells were plated out on thiamine a second time, viability of *cdc17*Δ cells harbouring nuclear Cdc17 without the PCNA binding motif was only slightly reduced compared to control cells (~ 80% versus >95%). Additionally, microscopic inspection revealed only marginally elongated cells. The findings of the follow-up experiments are illustrated in Figure 4.4.

In conclusion, if nuclear Cdc17 lacking the PCNA binding motif, co-expressed with a mitochondrial Cdc17, is subjected to conditions allowing only limited production of proteins, cell viability is reduced by 70-90%. Cell elongation implies that the DNA replication checkpoint machinery is called into action indicating that DNA replication is defective in these cells. However, prolonged growth under those conditions results in improved cell viability, with the viable cells dividing at a reduced cell size compared to initial plating on thiamine-containing media.

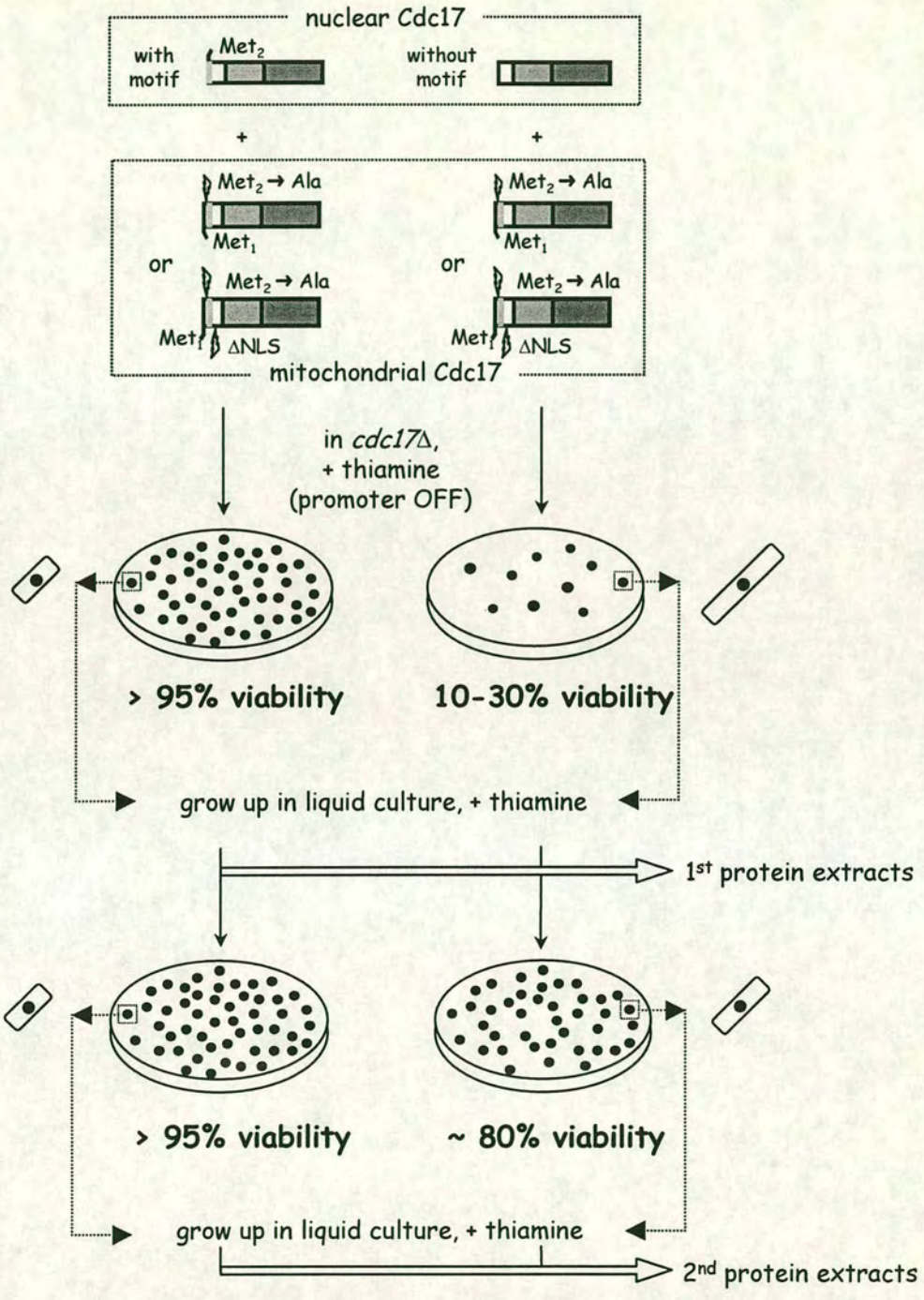



**Figure 4.4**: Summary of the results of the conducted double plasmid shuffle experiments. Nuclear proteins are Cdc17-NΔ19 (with motif) and Cdc17-NΔ31 (without motif) and the mitochondrial protein is either Cdc17-M20A or Cdc17-M20AΔNLS. See text for details.

### 4.3.4 Outlook

The most likely explanation of the observed behaviour of *cdc17*Δ cells expressing nuclear Cdc17 lacking the PCNA binding motif and a mitochondrial form, described in the previous section, is that cells with elevated Cdc17 levels are selected for. In order to test whether prolonged growth in thiamine-containing media causes increased Cdc17 levels, which would reflect such selective pressures, protein extracts were prepared from *cdc17*Δ cells of equally sized colonies which differ in the time-span they have been cultured on thiamine-containing media (see Figure 4.4). Preliminary immunoblotting analyses of these extracts with anti-Cdc17-antibody were performed. However, the resultant Western blots were difficult to interpret: many degradation products were detected and, in cells expressing more than one form of Cdc17, it was impossible to determine which protein species originated from which expression construct (data not shown). However, no significant increase in protein levels with respect to prolonged growth on thiamine was seen. It is worth noting, that the overall detected protein level was at least two-fold higher in cells expressing nuclear Cdc17 lacking the motif to those expressing nuclear Cdc17 with the motif suggesting that protein levels are a crucial factor in these experiments. In order to circumvent the issue of expression from multi-copy plasmids, it would be critical that nuclear Cdc17 was expressed from the chromosomal *cdc17* copy. Furthermore, to distinguish between nuclear and mitochondrial forms, differential tagging could be useful. Several attempts have been undertaken to C-terminally-tag chromosomal *cdc17* employing the PCR-mediated gene targeting method (Bähler *et al.*, 1998), but the tagged protein either could not be detected (GFP-tag, data not shown) or the presence of the tag impaired the growth properties of the strain (13myc-tag, data not shown).

## 4.4 Discussion

### 4.4.1 Overexpression of proteins with a PCNA binding motif

In section 4.2 the impact of high levels of Cdc17 proteins containing the PCNA binding motif on cell cycle progression was investigated. Expressing full-length Cdc17 or the presumed nuclear Cdc17-NΔ19 from the derepressed *nmt1* promoter (highest possible expression) caused Sp5 cells to elongate considerably, whereas gross overproduction of Cdc17 lacking the PCNA binding motif, Cdc17-NΔ31, only showed a minor cell cycle delay. These results fit in with previous studies overexpressing proteins containing $p21^{CIP1}$-like PCNA binding motifs in *S. pombe*. Heterologous expression of human p21 from the derepressed *nmt41* promoter in *S. pombe* caused an inhibition of cell cycle progress (Tournier *et al.*, 1996). In contrast, expression of a p21 protein with a mutation in the PCNA binding motif that is defective in PCNA binding abolished cell cycle inhibitory effects suggesting that the cell cycle delay was due to binding to PCNA (Tournier *et al.*, 1996). The authors also demonstrated that the degree of cell cycle inhibition correlated with p21 expression levels (Tournier *et al.*, 1996). Similarly, a C-terminal portion of Cdc27 including the PCNA binding site was found to cause significant cell elongation whereas a similar fragment lacking the motif did not (Reynolds *et al.*, 2000).

Interestingly, there are differences in the extent of cell cycle inhibition due to overexpression of motif-containing proteins which suggests that additional factors influence the ability of those proteins to bind to PCNA. The expression of a Cdc17 protein where amino acids 1-31 were replaced with the PCNA binding motif of *S. pombe* Rfc1 (amino acids 1-10) from the derepressed *nmt41* promoter was sufficient to exert a pronounced *cdc* phenotype. On the contrary, $cdc17^{+}$ expressed from the same promoter resulted in only a very minor cell cycle delay. It is unclear where the difference in cell cycle inhibitory effects originates from. As far as can be assessed from Western blotting analyses the two proteins are expressed to similar levels (data not shown). Comparison of the PCNA binding site regions reveals subtle differences in their amino acid composition, which may account for their differing binding

affinities to PCNA. Figure 4.6 shows an alignment of the respective PCNA binding motifs.

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Sp*Cdc17 | 20 | M | S | T | R | Q | S | D | I | S | N | F | F |
| *Sp*Rfc1 | 1 | | | M | S | N | S | D | I | R | S | F | F |

**Figure 4.5**: An alignment of PCNA binding motifs present in the Cdc17 protein which have been overproduced in *S. pombe*. The boxed residues are the conserved features of the motif. See text for details.

### 4.4.2 Co-expression of nuclear and mitochondrial Cdc17 in *cdc17*Δ

Experiments described in chapter 2 led to the conclusion that production of Cdc17 from the repressed *nmt41* promoter most closely reflected endogenous levels of the DNA ligase I protein. Studies presented in Chapter 3 elucidated that Cdc17 proteins are present in the nucleus and in mitochondria and identified forms of Cdc17 which are likely to be associated with only one cellular compartment. Functional complementation experiments using *cdc17*Δ revealed that Cdc17 function is required in both cellular compartments. The simultaneous expression of nuclear and mitochondrial Cdc17 resulted in rescue of a *cdc17* deletion strain. This was taken as an indication that nuclear and mitochondrial functions can be separated and that the expression of separate proteins ensures proper DNA ligase I activities.

The expression in *cdc17*Δ of a Cdc17 protein truncated for the first 31 amino acids – including the PCNA binding motif – alongside a mitochondrially targeted protein (from the repressed *nmt* promoter in pREP41X and pREP42X, respectively) resulted initially in a significant decrease in cell viability. The colonies that arose were few in number and varied considerably in size. The degree of cell elongation was roughly inversely proportional to the colony size. Whereas cells taken from medium to large colonies could be re-cultivated in the presence of thiamine, cell material from small colonies was unable to propagate any further. This suggests that the intracellular Cdc17 protein levels may play a crucial role for the survival of *cdc17*Δ cells. If

nuclear Cdc17 levels (lacking the motif) pass a threshold, then reasonably sized colonies can be formed, although the cells are delayed in cell cycle progression. However, if this threshold is not reached then small colonies are formed which eventually cease to grow.

Several parameters concerning the experimental strategy do not allow an unambiguous conclusion as to whether the PCNA binding motif is required for nuclear Cdc17 function. It is clearly not absolutely essential since experiments presented in Chapter 2 showed that 100 – 200-fold overproduction of a protein lacking the N-terminal 125 amino acids rescued *cdc17*Δ. This implies that high levels of Cdc17, truncated for the N-terminal domain can bypass the requirement for sub-cellular targeting signals and the presence of the PCNA binding motif. In conclusion, protein levels play a crucial role in assessing the importance of PCNA binding to Cdc17. Expressing proteins from multi-copy plasmids consequently results in considerable variation of intracellular mRNA and protein levels. Efforts aimed at providing uniform expression levels through expression from the chromosome have proven unsuccessful since it was impossible to tag chromosomal *cdc17*$^{+}$. Despite the described obstacles, it should be stressed, that expression of nuclear Cdc17 lacking the PCNA binding motif (alongside mitochondrial Cdc17) in a *cdc17*Δ strain, when produced at levels similar to those in wild-type cells, results in greatly reduced cell viability, compared to Cdc17 harbouring the motif. This, together with the overexpression studies, strongly suggests that the DNA ligase I – PCNA interaction serves a vital role in *S. pombe*.

## 4.5 Summary

The PCNA binding motif, which is present in the N-terminal domain of Cdc17, serves important functions in *S. pombe*. Overexpression of Cdc17 proteins harbouring its own motif or that of Rfc1 in wild-type cells causes a considerable cell cycle delay. Nuclear Cdc17, co-expressed with mitochondrial Cdc17, is able to rescue *cdc17*Δ. The absence of the PCNA binding motif from nuclear Cdc17 causes a severe reduction in cell viability.

# 5. Discussion

## 5.1 Summary of results

This section summarises the results that have been presented in this work.

The lesion sites in the *cdc17* temperature-sensitive mutants, that were part of the original *cdc* mutant collection, were determined. In each case a single nucleotide change resulted in a single amino acid change. In *cdc17-K42*, amino acid Glu298 is replaced by a lysine residue, in *cdc17-L16* Pro584 is converted to a serine and *cdc17-M75* harbours a mutation resulting in a Pro679 to serine change.

In a wild-type *S. pombe* strain, Cdc17 protein levels are comparable to those generated by expression of $cdc17^{+}$ from the repressed promoter in pREP41X.

Lack of the N-terminal 125 amino acids of Cdc17 fails to complement a *cdc17* deletion strain when expressed at levels similar to wild-type, whereas at a 100 – 200-fold overproduction rescue does occur. In contrast, expression of *cdc17* constructs encoding only the catalytic domains do not rescue *cdc17*Δ irrespective of the expression level.

The N-terminal domain of Cdc17 (amino acids 1-175) localises GFP to the nucleus and mitochondria. Removal of the first 19 amino acids abolishes mitochondrial targeting and alteration of amino acid Met20 to an alanine residue dramatically reduces nuclear localisation. Residues 62 Lys-Arg-Lys 64 are crucial residues for the nuclear localisation signal.

Nuclear and mitochondrial Cdc17 functions are both essential for *S. pombe* cells, since *cdc17*Δ is not complemented by proteins which are defective in localisation to either cellular compartment. It is possible to separate nuclear and mitochondrial functions of Cdc17, as judged by functional complementation of *cdc17*Δ by their co-expression. Lack of the PCNA binding motif from nuclear Cdc17 drastically reduces cell viability when expressed, together with mitochondrial Cdc17, at levels comparable to wild-type. Rescued cells are highly elongated, indicating a defect in DNA replication.

## 5.2 Introduction

In the following three sections published work that is related to this thesis is discussed and, where appropriate, compared to the results presented here. Section 5.3 concentrates on complementation studies of various DNA ligase proteins in different organisms. In section 5.4, mitochondrial DNA ligase functions are addressed as well as possible mechanisms for the dual targeting. Section 5.5 deals specifically with the interaction between DNA ligase I and PCNA and what purpose it may serve and in section 5.6 a brief summary of what is known about possible other DNA ligase I-interacting-proteins is given.

## 5.3 Structure-function analyses

Functional complementation experiments carried out in this work were aimed at clarifying the functions of the non-catalytic domains in DNA ligase I homologues.

DNA ligases all have versions of the catalytic domains in common, which make up the complete ATP-dependent enzyme of bacteriophages and are considered to be the minimal architecture necessary for carrying out the ligation reaction.

In the literature, several (cross-)complementing experiments have been described. For example, mutational analyses have been performed on human DNA ligase I cDNA expressing N-terminally truncated ligase proteins in a conditional-lethal $NAD^+$-dependent DNA ligase mutant of *E. coli*. Human DNA ligase I lacking the N-terminal 249 amino acids was capable of rescuing the temperature-sensitive phenotype whereas truncation of a further 35 amino acids could not complement. It has to be noted however that the former protein was expressed to a higher degree than the latter (Kodama *et al.*, 1991). This suggests that those 35 amino acids may contribute to the stability of the enzyme (see below).

Other studies have shown, that a human DNA ligase I truncation lacking the N-terminal 231 amino acids was unable to complement mouse embryonic stem cells with a lethal DNA ligase I null mutation. Even isolates where the truncation was expressed several fold higher than the wild-type gene in murine cells, also failed to rescue. In contrast, the same construct *was* able to complement the temperature-

dependent lethality of the *S. cerevisiae* ligase mutant *cdc9-7* (Petrini *et al.*, 1995). There are several possibilities as to why the N-terminal truncation does not rescue murine knock-out cells.

Firstly, it is conceivable that the quantity of truncated protein is not high enough for complementation of murine DNA ligase I knockout cells. Consistent with this, an N-terminal truncation of the *S. pombe* DNA ligase I homologue Cdc17, when expressed at levels similar to wild-type, is unable to rescue a deletion strain. However, 100 – 200-fold overexpression of this construct *is* able to support growth, indicating that high levels of truncated DNA ligase I protein can bypass the requirement for the N-terminal domain (Chapter 2, see also discussion section 5.5). The requirements to rescue a ligase-deficient yeast strain may be less demanding. For example, the level of the ectopically expressed truncation (under the control of the strong *ADH1* promoter) in yeast relative to wild-type Cdc9p is not known, but is likely to be high. Although the activity of the truncated human DNA ligase I protein *was* sufficient to permit growth of *cdc9-7*, the ligase activity in this case was less than 20% compared to wild-type Cdc9p (Petrini *et al.*, 1995), supporting the conclusion that full rescue cannot be achieved by an N-terminal truncation.

Secondly, N-terminally truncated DNA ligase I proteins may lack species-specific interaction sites for regulatory and/or replication proteins which are essential for proper functioning at approximately endogenous levels. The N-terminus, for example, binds to PCNA, an interaction which will be discussed in section 5.5 in more detail.

As yet, no function has been assigned to the conserved non-catalytic domain (CNCD). Several lines of evidence suggest that the CNCD contributes to an increased stability of the enzyme. Human DNA ligase I starting at residue 285 expressed in *E. coli* was produced to a much lower extent than truncated proteins starting at amino acid 220 or 250 (Kodama *et al.*, 1991). Experiments presented in this work show that the catalytic core without the CNCD can only be produced to a maximum of 25% compared to the wild-type protein according to immunoblotting analyses (section 2.5).

Human DNA ligase III and IV are both found as heterodimers in interaction with proteins that stabilise the ligase (XRCC1 and XRCC4, respectively). Strains deficient for these accessory proteins display reduced levels of the respective ligases. These associations occur through the C-terminal extensions of the ligases, which contain putative protein-interacting domains. In the case of DNA ligase IV, the association with an XRCC4 homologue is conserved in yeast. DNA ligase I does not contain any C-terminal modules but instead it has an extended middle domain (in addition to the proteolytically sensitive N-terminal domain discussed above), which is conserved between cellular replicative DNA ligase homologues. It is conceivable that this middle domain serves to stabilise the ligase, possibly via an interaction with an (as yet) unidentified protein partner. Perhaps this partner may even be ligase itself since biophysical experiments suggested that DNA ligase I is able to trimerise (Dimitriadis *et al.*, 1998). However, this association was dependent on salt concentration and temperature and the fact that a homo-oligomerisation has not been described elsewhere raises the question as to whether or not this is biologically significant.
The CNCD may also well be the target of interaction sites of DNA replication/repair proteins which anchor the protein in the replisome/repairosome. Alternatively – or additionally – this domain may be responsible for increasing the efficiency of ligation. A recent publication described the finding that the eukaryotic single-stranded binding protein RPA was able to stimulate *in vitro* DNA ligase I activity by specifically increasing the rate of catalysis (Ranalli *et al.*, 2001). Since T4 DNA ligase (which lacks a CNCD) was not affected, the stimulation may well be mediated via the middle domain. Whether this has any functional implications for the *in vivo* situation remains to be determined.

In addition to the complementation of a temperature-sensitive *E. coli* ligase mutant by an N-terminally truncated version of human DNA ligase I (see above), other reports also describe the complementation of ligase-deficient strains with enzymes which are only distantly related on an evolutionary basis. For instance, *Salmonella typhimurium* with a disrupted $NAD^+$-ligase could be apparently fully complemented by the ATP-dependent T4 DNA ligase (Park *et al.*, 1989). It appears that an $NAD^+$-

ligase function can be substituted for by their ATP-dependent counterparts, probably when present at high levels.

Rather surprisingly, there is a recent report which describes the apparently full complementation of *S. cerevisiae cdc9*Δ and *cdc9*Δ *lig4*Δ strains by expression of the ATP-dependent *Chlorella* virus ligase (only consisting of the catalytic core domains) or by the $NAD^+$-ligase of *E. coli* (Sriskanda *et al.*, 1999). It should be noted, however, that while wild-type *CDC9* was expressed under its own promoter the heterologous ligases were expressed under the control of the strong *TPI1* promoter from *CEN* plasmids. Triosephosphate-isomerase is a key enzyme in glycolysis, the first step in the major energy metabolic pathway, with a highly active promoter. Thus, it is possible that proteins are expressed at very high levels from this promoter. Further, the researchers found that high levels (expressed from the *TPI1* promoter) of Cdc9 proteins with deletions of the 178-343 N-terminal amino acid residues could not rescue *cdc9*Δ or *cdc9*Δ *lig4*Δ supporting the proposal that the middle domain may increase enzyme stability. Consistent with this finding, experiments with the fission yeast DNA ligase I homologue demonstrated that a Cdc17 protein deleted for the N-terminal 390 amino acids and only consisting of the catalytic domains was unable to rescue a deletion strain even at high levels of expression (Chapter 2). It appears that while high levels of the distantly related *Chlorella* virus ligase can rescue a deletion strain, the homologous portion of the endogenous enzyme cannot. In this regard it may be interesting to establish if the additional motifs present in the C-terminal portion of $NAD^+$-ligases, namely a zinc finger, HhH motif and BRCT domain (Doherty and Suh, 2000), are required for rescue of eukaryotic species devoid of an ATP-ligase.

It should be stressed that complementation with heterologous DNA ligase proteins requires their gross overproduction, which does not reflect physiological conditions.

It is proposed that the conserved non-catalytic domain serves to stabilise the enzyme and that factors intrinsic to this domain may additionally increase the overall efficiency of the enzyme. With regard to the rescue of an *S. pombe* strain deleted for *cdc17*, the presence of the CNCD was required under all tested conditions. Hence, the conserved non-catalytic domain performs essential functions *in vivo*.

The N-terminal domain of the *S. pombe* DNA ligase I protein Cdc17 performs important functions *in vivo*, which will be discussed in greater detail in the following two sections, but overexpression of the rest of the enzyme can circumvent its requirement.

## 5.4 DNA ligase in mitochondria

### 5.4.1 Function in DNA repair

In this work it has been shown that the *S. pombe* DNA ligase I homologue Cdc17 is localised in the mitochondria as well as the nucleus. Previous work demonstrated the presence of a mitochondrial DNA ligase in *Xenopus* and humans. In these cases the ligase proteins originate from the *LIG3* gene, the nuclear product of which is thought to be involved in the final step of base-excision repair (BER). Recently, in *S. cerevisiae* the existence of a mitochondrial Cdc9 protein was reported. Cdc9p is a DNA ligase I homologue and constitutes the replicative ligase function during chromosomal DNA replication in budding yeast. *S. cerevisiae* and *S. pombe* genomes do not have genes encoding DNA ligase III homologues.

In the past few years, several other components known to be involved in the nuclear BER pathway have been identified in human mitochondria, that is an AP endonuclease encoded by *APE2* and an *UNG* encoded uracil-glycosylase (see Introduction). Pinz and Bogenhagen, 1998, demonstrated that *Xenopus* mitochondria possess all the factors necessary to perform BER, including the mitochondrial DNA polymerase pol γ and a ligase (for a recent review on mitochondrial BER see Bogenhagen *et al.*, 2001). The existence of the BER-like UVER pathway in *S. pombe* indicates that DNA repair of some types of lesions, namely those that can be generated endogenously through reactive oxygen species, is conserved in mitochondria from yeast to man.

### 5.4.2 Function in DNA replication

Although great advances have recently been made on DNA repair in mitochondria, much less is known about replication of mitochondrial DNA. Early work suggested that mammalian mtDNA is replicated by a strand-asymmetric mechanism, where leading strand (heavy- or H-strand) synthesis is initiated from its origin and once the lagging strand (light- or L-strand) origin is displaced, synthesis of the latter strand begins in the opposite direction (Clayton, 1982, reviewed in Clayton, 1996). Recent analyses of mammalian mtDNA replication intermediates by 2D-gel-electrophoresis identified replication products which most likely arise from coupled leading- and lagging-strand synthesis, in addition to products presumed to derive from the strand-asynchronous-mode of DNA replication (Holt *et al.*, 2000). Application of the same technique to the analysis of *S. pombe* mtDNA replication yielded products which prompted the authors to suggest a rolling-circle-type model of DNA replication (Han and Stachow, 1994). Irrespective of the precise mechanism, the fact that human and yeast mtDNA generally seem to adopt the conformation of covalently closed circles, dictates the necessity for the participation of a DNA ligase during mtDNA replication.

Expression of mitochondrial genes serves an essential function in most organisms and a prerequisite of this is an intact mitochondrial genome. Since a lack of mitochondrial Cdc9 in *S. cerevisiae* or diminished levels of mitochondrial DNA ligase III in human HT1080 cells has been shown to result in a reduction of mtDNA content, it can be concluded that the respective ligase proteins perform vital functions in mtDNA replication (Donahue *et al.*, 2001, Lakshmipathy and Campbell, 2001). Experiments presented in this thesis underscore the importance of mitochondrial DNA ligase function: the absence of mitochondrial Cdc17 in *S. pombe* rendered cells inviable. In this regard it is worth noting that *S. pombe* is considered a petite-negative yeast. This is in contrast to the petite-positive *S. cerevisiae*. Budding yeast is able to tolerate large deletions in or the complete absence of mitochondrial DNA ($\rho^-$ or $\rho^0$ strains, respectively), the result of which is the formation of small (= petite) colonies when grown on non-fermentable media. In other words, the survival of *S. pombe* cells depends on the expression of mitochondrial genes.

Since the ligase I genes in yeasts and ligase III genes in *Xenopus* and humans appear to encode the only mitochondrial ligase, the respective ligase proteins seem to unify mitochondrial DNA replication and repair functions.

### 5.4.3 Dual localisation of Cdc17

*S. pombe cdc17* belongs to a class of genes which encode isozymes that are found in the nucleus as well as in mitochondria. Characteristic for genes encoding such proteins is the presence of two (or more) in-frame translation-initiation codons in their 5' region. In section 1.4 it was described that the usage of alternative start codons yielding proteins that possess or lack a mitochondrial targeting sequence (MTS) can occur either at the transcriptional or the translational level. Figure 5.1 gives an outline of how *cdc17* may encode nuclear and mitochondrial versions of DNA ligase.

When *cdc17* was originally characterised, S1-nuclease-mapping revealed the existence of a major and minor transcription initiation start site as judged by the relative abundance of two mRNA species (Barker *et al.*, 1987). The predicted 5' termini of the transcripts would correlate with translation of a predominantly nuclear Cdc17 protein from the major transcript ($cdc17_{nucl}$-mRNA), where translation starts at the second AUG codon of the initially identified ORF (which represents the first accessible in this message), and the production of a less copious mitochondrial version from the minor transcript ($cdc17_{mito}$-mRNA). This presents an elegant solution to the different requirements of ligase activity according to the varying quantity of DNA present in the organelles: considering that the copy number of the mitochondrial genome (~19 kb) per cell is several hundred at most (Bostock, 1969) the overall DNA content in mitochondria is on average 10% of the nuclear genome.

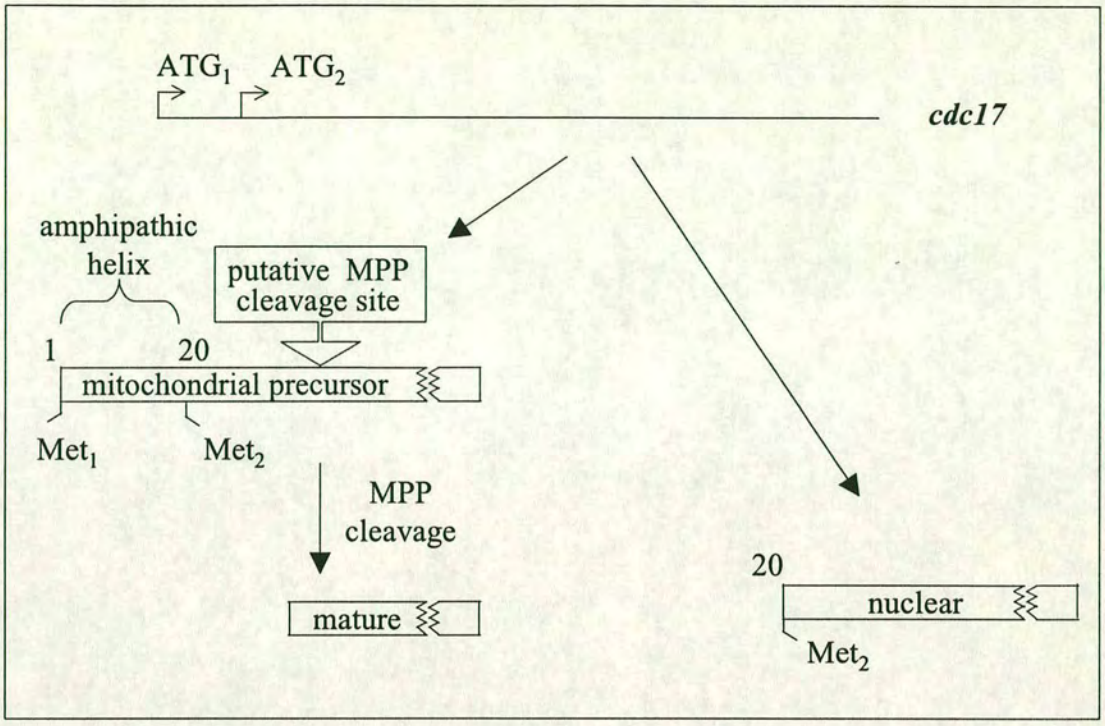



**Figure 5.1**: Model of how the *cdc17* gene can give rise to two proteins targeted to mitochondria and the nucleus. Note that either two transcripts can be produced which each are translated or alternative translation initiation from one mRNA is used (see Chapter 1). MPP stands for mitochondrial processing peptidase which cleaves off the presequence after the mitochondrial import machinery has recognised the N-terminal signal peptide. The precise location of the MPP cleavage site is not known.

The only targeting information known to be present in the presumptive nuclear Cdc17 protein is the nuclear localisation signal (NLS), which is partly represented by amino acids 64 KRK 66 (Chapter 3). The mitochondrial Cdc17 protein precursor, however, harbours the upstream MTS in addition to the NLS. What mechanisms ensure the mitochondrial targeting of this translated protein in preference to its nuclear localisation? One possible explanation takes the kinetics of translation and translocation inside a cell into consideration. If a translationally active mRNA, encoding a mitochondrial precursor, happens to be in proximity to the mitochondrial import machinery, it is likely that subsequent preprotein molecules are imported co-translationally. In keeping with this model, electron microscopy studies have shown that translationally active ribosomes accumulate on the surface of yeast mitochondria (reviewed in Lithgow, 2000). Alternative or additional possibilities include that translation initiation from $cdc17_{mito}$-mRNA starts at both the upstream and downstream start codon and/or that a proportion of Cdc17 proteins harbouring the MTS also localises to the nucleus. The latter suggestion would account for the observation that a GFP-fusion of the N-terminus of Cdc17, in which the second start methionine was replaced by an alanine, showed weak nuclear fluorescence. However, this observation could also be due to high levels of $cdc17_{mito}$-mRNA, expressed from pREP3X, leading to the mitochondrial protein import machineries working to capacity. Attempts to perform experiments to identify which forms of Cdc17 localised to which compartment were hampered by the fact that C-terminal tagging of *cdc17*$^{+}$ impaired growth. Also, sub-cellular fractionation could not be monitored satisfactorily.

## 5.5 DNA ligase I – PCNA interaction

DNA ligase I and its homologous proteins harbour a $p21^{CIP1}$-like PCNA binding motif in the extreme N-terminus. The N-terminal domain of the *S. pombe* DNA ligase I protein Cdc17 binds the fission yeast PCNA-homologue Pcn1 *in vitro* and as part of this work it was shown that rescue of *cdc17*Δ with a nuclear Cdc17 protein that is lacking the motif dramatically reduced cell viability, when expressed at

physiological levels. Surviving cells displayed a *cdc* phenotype indicative of defects in S phase.

Although previous studies have already shown that the presence of the conserved motif in human DNA ligase I is necessary for localisation to sub-nuclear foci (Montecucco *et al.*, 1998), the work presented here further supports the conclusion that the association between ligase and PCNA is, under physiological conditions, required for *S. pombe* cell survival. This requirement can however be bypassed by high levels of N-terminally truncated ligase protein. This observation is in accordance with several published reports. The *in vitro* SV40 DNA replication system has been used to test the ability of human 46BR.1G1 cell extracts, which possess a mutationally altered DNA ligase I, to perform complete DNA synthesis. Although 46BR.1G1 cells have normal ligase protein levels, they only display 5% of wild-type DNA joining activity. Consistent with this, DNA synthesis by these cell extracts was incomplete, leaving behind Okazaki fragments that were not ligated together. Addition of DNA ligase I corrected this defect. An equivalent amount of enzyme lacking the N-terminal domain could not correct aberrant DNA replication, whereas addition of a ten-fold excess of this truncation supported efficient DNA synthesis (Prigent *et al.*, 1994, Mackenney *et al.*, 1997). Furthermore, the same DNA replication system (*in vitro* SV40 DNA replication by 46BR.1G1 cell extracts) was used by Levin *et al.*, 2000. They showed that the full-length DNA ligase I, where the adjacent phenylalanine residues of the PCNA binding motif had been replaced by alanines, did not complement the replication defect when an equivalent amount of wild-type protein corrected the deficiency. Again, addition of a ten-fold excess of the mutant DNA ligase I protein resulted in completely ligated DNA replication products (Levin *et al.*, 2000). Although the above described experiments were carried out *in vitro*, experiments in this work suggest that approximately physiological amounts of nuclear *S. pombe* DNA ligase I without the motif were insufficient to restore normal DNA replication *in vivo* since a deletion strain could not be functionally complemented.

What is the precise function of DNA ligase I binding to PCNA? The results of several studies argue against the possibility that PCNA directly stimulates catalytic activity of DNA ligase I. Preincubation of DNA substrates containing a single nick

with PCNA did not have a significant effect on DNA joining upon addition of DNA ligase I (Levin *et al.*, 1997). Although Tom *et al.*, 2001 did find a stimulation of ligase acitivity, they showed that it was coupled to PCNA being bound on the nick-containing-DNA. Furthermore, the fact that PCNA incubated with T4 DNA ligase did not show any stimulatory effect on ligation activity most probably rules out a direct influence on catalytic activity. It is more likely that the presence of PCNA on DNA increases the rate of DNA ligase I-binding to the substrate via an interaction with PCNA. This would explain that, at physiological levels of DNA ligase I, the PCNA-interaction-site is required to target DNA ligase I to nicks, whereas high levels of DNA ligase protein can circumvent the "substrate-targeting-sequence", because statistically it is very probable that a ligase molecule will be close to its action site. This explanation is consistent with the findings presented in this work and described above, and also with the observation that the PCNA binding motif is necessary for specific localisation to sub-nuclear replication foci (Montecucco *et al.*, 1998).

## 5.6 Other DNA ligase I interacting proteins

In the literature physical interactions of DNA ligase I with two proteins apart from PCNA have been described so far. Interestingly, both proteins are not thought to be involved in DNA replication, but are rather components of DNA repair pathways. These proteins are DNA polymerase β, a polymerase functioning in base-excision repair (Prasad *et al.*, 1996) and Mre11, which is implicated in double-strand break repair, specifically meiotic recombination and non-homologous-end-joining (Petrini *et al.*, 1995b). The latter finding is rather surprising since specifically DNA ligase IV homologues have been implied in functioning in double-strand repair pathways.

Therefore, so far it appears as if DNA ligase I only interacts with the DNA replication protein PCNA (which is also involved in DNA repair pathways). In the original report describing an interaction between DNA ligase I and PCNA, PCNA was discovered as a protein from Hela cell nuclear extracts specifically retained on a DNA ligase I affinity resin (Levin *et al.*, 1997). Similarly, Mre11 was discovered as

an interactor in a two-hybrid-screen. To date, no other interacting proteins have been published although screens have been carried out. This could partly be due to the fact that the N-terminal domain of DNA ligase I proteins is highly sensitive to proteolytic degradation. Since attempts to C-terminal tag the *S. pombe* DNA ligase I gene *cdc17*$^{+}$ has proven unsuccessful (data not shown), Cdc17-associating proteins could not be identified by co-immunoprecipitation. In this lab, work is underway to produce the middle conserved non-catalytic domain of Cdc17, which will subsequently be tested for interactors (Leonora Ciufo, personal communication).

A correlation between DNA ligase I and other proteins involved in Okazaki fragment processing – other than PCNA – has so far only come from the detection of genetic interactions. In this regard, overexpression of *cdc17*$^{+}$ rescues the temperature-sensitive lethality of *dna2* temperature-sensitive mutants. Also, *cdc17-K42* is synthetically lethal with *dna2* mutants, which has also been shown to be the case for the respective budding yeast proteins (Kang *et al.*, 2000, Ireland *et al.*, 2000).

As part of this work, *cdc17*$^{+}$ has been tested specifically for interactions with several DNA replication proteins using a two-hybrid assay. However, no interactions were observed between Cdc17 and itself, Rad2, Dna2, Cdc24, Cdc1 and Cdc27 using the overlay assay. Moreover, a two-hybrid screen was carried out using the conserved non-catalytic and catalytic domains of the budding yeast DNA ligase I homologue Cdc9 as a bait against a *S. cerevisiae* cDNA library. Again, no obviously meaningful interactors were recovered.

## 5.7 Conclusions

The *S. pombe cdc17*$^{+}$ gene, which encodes the DNA ligase I homologue, is essential for cell viability (Nasmyth, 1977 and this work). In addition to the C-terminal catalytic domains, the middle conserved non-catalytic domain is necessary for functional complementation of a *cdc17* deletion strain. The latter domain appears to be required for increased stability of the enzyme. The presence of the N-terminal domain of Cdc17 is important for complementing *cdc17* deficient strains at physiological levels. Intrinsic to this domain are sub-cellular localisation signals

signals, namely nuclear and mitochondrial targeting sequences, as well as a PCNA binding motif, which serves as targeting the ligase to nicked DNA.

No genetical or physical interactions with other DNA replication proteins were found; this does, however, not rule out their existence. It would be highly interesting to identify possible interactors of DNA ligase I at the replication fork, which may give further insights into the precise mechanism of the abundant process of Okazaki fragment maturation during DNA replication. In this regard, it would also be worthwhile to determine any DNA ligase-associated proteins in mitochondria, in order to aid understanding of mitochondrial DNA replication and repair.

# 6. Materials and Methods

## 6.1 General

Many of the standard molecular biological and biochemical methods were carried out on the basis of the protocols described in Sambrook *et al.*, 1989.

### 6.1.1 Chemicals

Chemicals were purchased from Amersham, BDH, Fisher, Fluka or Sigma unless otherwise stated.

### 6.1.2 Enzymes

Restriction enzymes, DNA polymerases and other enzymes used in this work were purchased from Gibco BRL, NEB, Promega, Roche and Sigma unless otherwise stated.

### 6.1.3 Commonly used reagents

| | |
|---|---|
| CPS | 50 mM citrate-phosphate buffer, pH 5.6, 1.2 M sorbitol, 0.1 % (v/v) β-mercaptoethanol, 40 mM EDTA |
| DNA loading buffer (6x) | 0.25% bromophenol blue, 0.25% xylene cyanol FF, 40% sucrose |
| LiAc (10x) | 1 M lithium acetate, pH 7.5 (dil. acetic acid) |
| PBS (10x) | 115 g/l $Na_2HPO_4$, 29.6 g/l $NaH_2PO_4$, 58.4 g/l NaCl |
| PBST | 1x PBS, 0.1% (v/v) Tween-20 |
| Protein gel running buffer | 3 g/l Tris, 14.4 g/l glycine, 1 g/l SDS |
| SDS sample buffer (2x) | 100mM Tris-HCl, pH 6.8, 200mM DTT (added fresh immediately before use), 4% (w/v) SDS, 0.2% (w/v) bromophenol blue |

| | |
|---|---|
| 1M sodium phosphate buffer, pH 7.0 | 57.7 ml/l 1M $Na_2HPO_4$, 42.3 ml/l 1M $NaH_2PO_4$ |
| TAE (50x) | 242 g/l Tris, 57.1 ml/l glacial acetic acid, 100 ml/l 0.5 M EDTA (pH 8.0) |
| TE (10x), pH 7.5 | 100 mM Tris-HCl, 10 mM EDTA, pH 7.5 |
| TE (10x), pH 8.0 | 100 mM Tris-HCl, 10 mM EDTA, pH 8.0 |
| Transfer buffer | 14.4 g/l glycine, 3 g/l Tris, 20% methanol |
| TSB | 10% PEG 3000, 10 mM $MgCl_2$, 10 mM $MgSO_4$ 5% DMSO in LB |
| TSBG | TSB + 20 mM glucose |

## 6.2 Oligonucleotides

Oligonucleotides were routinely ordered from Amersham Pharmacia or MWG Biotech. Stock solutions were 100 µM and working concentrations were 10 µM for PCR amplification and 1.6 µM for sequencing reactions. The oligonucleotides used in this study are listed below. The positions within *cdc17* are numbered according to Appendix 1. Restriction sites are underlined. F refers to forward primer and R to reverse primer. Oligonucleotides used for site-directed mutageneses are listed in section 6.11.

| name | description | sequence |
|---|---|---|
| #1 | cloning of *cdc17* by PCR, F, introduction of *BamH*I site | GATCGGATCCTTGTTATGCGAACAGTATTT TCG |
| #2 | cloning of *cdc17* by PCR, R, introduction of *BamH*I site | GATCGGATCCCAAACTTAATAATCTTCGGC AGCTGGGGAC |
| #3 | for *cdc17* deletion, F, 1141-1219F + 20 bp of pFA6a (upstream of *kanMX6* module) (Bähler *et al.*, 1998) | CGTACACGTGTTACTTGTTCTTCAGGTGTAC CCTCACTATTGTCAAATCTTATGAAATTCTT TACTCGTTGACTTGATTCGGATCCCCGGGT TAATTAA |
| #4 | for *cdc17* deletion, R, 3719-3637R + 20 bp of pFA6a (downstream of | ATTCATATTAGCCTCTTTGGGCAAATTTTAG CCGAACATTACCCTTTATTATGACAGTAGC |

| | | |
|---|---|---|
| | *kanMX6* module) | TCTCAGCAGTAACTCTCAGAATTCGAGCTC |
| | (Bähler *et al.*, 1998) | GTTTAAAC |
| #5 | verification of *cdc17* deletion, *cdc17* 5' sequence, 1006-1035F | GAGACAAGGAACACTCATACTATATCAATG |
| #6 | verification of *cdc17* deletion, *kanMX6* module sequence F | GCTAGGATACAGTTCTCACATCACATCCG |
| #7 | verification of cdc17 deletion, *cdc17* 3' sequence, 3769-3750R | AATAGCCATCAGCTAGAGCA |
| #8 | verification of *cdc17* deletion, *kanMX6* module sequence R | GATCGCAGTGGTGAGTAACCATGCATCATC |
| #0Seq | *cdc17* sequence, 1511-1492R | TCCAGTGTCGTCGTGGTTGT |
| #1Seq | *cdc17* sequence, 1622-1641F | GGCATCTACTCCAATACCT |
| #2Seq | *cdc17* sequence, 2033-2052F | GGTGATCTTGGTCTTGTACC |
| #3Seq | *cdc17* sequence, 2427-2446F | TGATTGAGCATGGACTTGGT |
| #4Seq | *cdc17* sequence, 2816-2835F | GACATTAAGGTCCGAGCATG |
| #5Seq | *cdc17* sequence, 3364-3383F | GCATCAGCCTGACGTCTGGT |
| #6Seq | *cdc17* sequence, 1713-1695R | CAGCAAATGTTGCATGACC |
| #7Seq | *cdc17* sequence, 2671-2652R | GATGTCCGGATAGCGAACAG |
| #8Seq | *cdc17* sequence, 2197-2177R | CAATAGCCGCTTGATCACTCC |
| #9Seq | *cdc17* sequence, 3602-3623F | GAAGGGTCCCCAGCTGCCGAAG |
| #pTZH | pTZ19R forward primer | GTAAAACGACGGCCAGT |
| #pTZE | pTZ19R reverse primer | TAATACGACTCACTATGG |
| #nmt5' | pREP3X forward primer | GGAATCCTGGCATATCATCAATTG |
| #nmt3' | pREP3X forward primer | GCAGCTTGAATGGGCTTCCATAGT |
| #MT1 | *S. pombe* mating type | AGAAGAGAGAGTAGTTGAAG |
| #MP | *S. pombe* mating type | ACGGTAGTCATCGGTCTTCC |
| #MM | *S. pombe* mating type | TACGTTCAGTAGACGTAGTG |

## 6.3 Plasmids

### 6.3.1 General

The following plasmids were used in this work.

| name | description | reference |
|---|---|---|
| pFA6a-kanMX6 | contains the *kanMX6* module, which was used to replace *cdc17* in order to generate *cdc17*Δ, used as template in PCR reaction to amplify *kanMX6* with suitable 5' and 3' ends (see section 6.6.10) | Bähler *et al.*, 1998 |
| pGEX6P-1B | *E. coli* expression vector | modified from Clontech |
| pGEX4T-cdc17 [1-175]-GFP | basis for cloning of Cdc17-GFP-fusions | Stuart MacNeill |
| pREP3X | *S. pombe* expression vector (section 6.3.2) | Maundrell, 1990, Forsburg, 1993 |
| pREP41X | *S. pombe* expression vector (section 6.3.2) | Basi *et al.*, 1993, Forsburg, 1993 |
| pREP81X | *S. pombe* expression vector (section 6.3.2) | Basi *et al.*, 1993, Forsburg, 1993 |
| pTZ19R | general purpose cloning vector | |
| pTZ19R-cdc17 | contains bp 1218-3636 of the 4176 bp *Hind*III-*Bgl*II fragment encoding full-length *cdc17*, cloned into *BamH*I site, used as basis for construction of mutants[1] | this work |
| pTZ19R-cdc17[1-4176] | contains 4176 bp *Hind*III-*Bgl*II fragment encompassing the entire *cdc17* ORF | Stuart MacNeill |
| pUR19-cdc17 | recovered from library as rescuer of *cdc17-K42*, map not known | Barbet *et al.*, 1992 |

[1] A DNA fragment carrying the *cdc17* gene was amplified by PCR from the template pTZ19R-cdc17[1-4176] with primers #1 and #2. The PCR fragment was digested with *BamH*I and cloned into the *BamH*I site of pTZ19R. The resultant pTZ19R-cdc17 was sequenced with primers #pTZH, #pTZE, #OSeq, #1Seq, #2Seq, #3Seq, #4Seq and #5Seq to confirm an error-free *cdc17* copy.

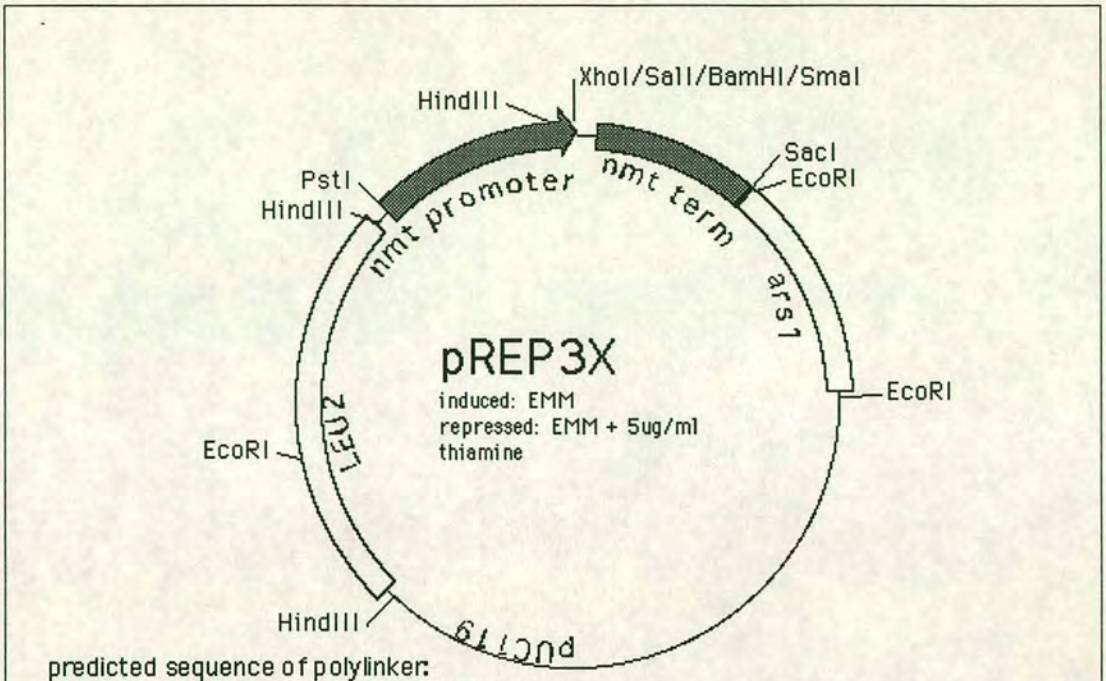
## 6.3.2 pREPX plasmids

### 6.3.2.1 Introduction

The pREP plasmids are the most commonly used *S. pombe* expression vectors and they contain the thiamine-regulatable *nmt* promoter (Maundrell, 1990). There are three versions of the *nmt* promoter: in addition to the original, two attenuated versions with reduced activity have been constructed through mutations in the TATA box. The respective plasmids are called pREP41 and pREP81 (Basi *et al.*, 1993). The original polylinker sequence containing an ATG codon was altered, and the resulting plasmids were are from the pREPX series (Forsburg, 1993, see Figure 6.1). In this work the following designation is used:

| promoter | description | |
|---|---|---|
| *nmt1* | *nmt* promoter in pREP3X | full strength |
| *nmt41* | *nmt* promoter in pREP41X | medium strength |
| *nmt81* | *nmt* promoter in pREP81X | low strength |

The promoter are fully induced in the absence of thiamine and fully repressed when thiamine was added to media to a final concentration of 5 μg/ml (stock 5mg/ml in $H_2O$, store at room temperature). The expression levels differ approximately ten-fold between usage of different promoters and approximately 100-fold between induced and repressed conditions (Basi *et al.*, 1993)




predicted sequence of polylinker:
TGGCCTCGAGGTCGACTCTAGAGGATCCCCGGG

NOTE: Xba is not unique; there are two XbaI sites in the ars.
REPs 41X/42X/81X/82X have no Sal site in the polylinker;
predicted sequence is CATAC CTCGAGGGATCCCCGGG

REP3X is derived from original REP3 series by addition of a Xho linker between BalI and SalI; this deletes the ATG within the polylinker, destroys BalI, and recreates SalI. Weaker REPXs (41,81,etc) were constructed by putting the Xho linker between BamHI and NdeI; the NdeI, SalI anmd XbaI sites are lost but BamHI is recreated.

Derivatives of the REP family:
RIP3, RIP3X (no ars, LEU2); RIP4, RIP4X (no ars, ura4+)
REP41 and REP41X (weaker promoter, LEU2)
REP81 and REP81X (weakest promoter, LEU2)
REP42 and 42X (weaker, ura4+); REP82 and 82X (weakest, ura4+)
Polylinker for 41X/81X is Xho/Bam/Sma
Polylinker for 41/81 is Nde/Sal/Bam/Sma

Reporter constructs: REP3X-lacZ, REP41X-lacZ, REP81X-lacZ

References:
REP3, parent vectors Maundrell (1993) Gene 123:127
REP41/42/81/82: basi et al (1993) Gene 123:131
REPX constructs: Forsburg (1993) NAR 21:2955

**Figure 6.1**: Map of pREP3X. The polylinker sequence for pREP3X as well as for pREP41X and pREP81X is given. The construction of pREP3X (-41X, -81X) from the original pREP3 vector is described. Several related vectors are stated, among them pREP42X, which is identical to pREP41X except that it contains the *ura4*$^{+}$ marker gene instead of *LEU2,* and which is also used in this study. The diagram is taken from the Forsburg pombe page (pingu.salk.edu/~forsburg/images/REP3X.gif).

### 6.3.2.2 Constructs used in complementation experiments

Most of the following *cdc17* constructs were engineered from pTZ19R-cdc17 through site-directed mutagenesis or sub-cloning from appropriate constructs (see section 6.11). After sequencing of the (mutagenic) *cdc17* sequence, the *BamH*I fragment was subcloned into pREPX vectors. Positions of amino acids are as in Appendix 2.

| construct | encoding protein |
|---|---|
| pREP3X-cdc17 | full-length Cdc17 (amino acids 1-768) (high levels) |
| pREP41X-cdc17 | full-length Cdc17 (amino acids 1-768) (medium levels) |
| pREP81X-cdc17 | full-length Cdc17 (amino acids 1-768) (low levels) |
| pREP41X-cdc17-NΔ31 | Cdc17 lacking the N-terminal 31 amino acids |
| pREP41X-cdc17-NΔ65 | Cdc17 lacking the N-terminal 65 amino acids |
| pREP41X-cdc17-NΔ96 | Cdc17 lacking the N-terminal 96 amino acids |
| pREP41X-cdc17-NΔ125 | Cdc17 lacking the N-terminal 125 amino acids |
| pREP41X-cdc17-NΔ390 | Cdc17 lacking the N-terminal 390 amino acids |
| pREP41X-cdc17-CΔ378 | Cdc17 lacking the C-terminal 378 amino acids |
| pREP41X-cdc17-K416A | full-length Cdc17 with a point mutation encoding an alanine at position 416 instead of a lysine |
| pREP41X-cdc17-M20A | full-length Cdc17 with a point mutation encoding an alanine at position 20 instead of a methionine |
| pREP42X-cdc17-M20AΔNLS | full-length Cdc17 with point mutations encoding an alanine at position 20 instead of a methionine and alanines at positions 62-64 instead of Lys-Arg-Lys |
| pREP41X-cdc17-NΔ31[1-10rfc1] | Cdc17 lacking the N-terminal 31 amino acids, which are replaced with amino acids 1-10 of *S. pombe* Rfc1 |
| pREP3X-lpp1-CT | catalytic domains of *S. pombe* DNA ligase homologue Lpp1 (provided by Stuart MacNeill) |
| pREP41X-lpp1-CT | catalytic domains of *S. pombe* DNA ligase homologue Lpp1 (provided by Stuart MacNeill) |
| pREP3X-SsoLig | DNA ligase of the archaeon *S. solfataricus* (provided by Stuart MacNeill) |
| pREP41X-SsoLig | DNA ligase of the archaeon *S. solfataricus* (provided by Stuart MacNeill) |

### 6.3.2.3 Constructs used for localisation studies

The relevant *cdc17* sequence with an upstream *Xho*I site and a downstream *Eco*RI site, generated by PCR (see section 6.11), was cloned into pGEX4T-cdc17[1-175]-GFP, where the existing *cdc17* fragment had been excised with *Xho*I/*Eco*RI. Subsequently, the *Xho*I/*BamH*I fragment encompassing *cdc17*-GFP sequences were subcloned into pREPX vectors. The cloning strategy is summarised in Figure 6.2.

| construct | Cdc17 sequence N-terminal to GFP |
|---|---|
| pREP3X-GFP | --- (provided by Stuart MacNeill) |
| pREP3X-cdc17[1-175]-GFP | N-terminal 175 amino acids (provided by Stuart MacNeill) |
| pREP3X-cdc17[20-175]-GFP | amino acids 20-175 |
| pREP3X-cdc17[32-175]-GFP | amino acids 32-175 |
| pREP3X-cdc17[66-175]-GFP | amino acids 66-175 |
| pREP3X-cdc17[97-175]-GFP | amino acids 97-175 |
| pREP3X-cdc17[126-175]-GFP | amino acids 126-175 |
| pREP3X-cdc17[1-175]M20A-GFP | amino acids 1-175, but alanine replacing methionine at position 20 |
| pREP3X-cdc17[1-175]M20AΔNLS-GFP | amino acids 1-175, but alanines replacing methionine at position 20 and Lys-Arg-Lys at positions 62-64 |
| pREP3X-cdc17[1-175]ΔNLS-GFP | amino acids 1-175, but alanines replacing Lys-Arg-Lys at positions 62-64 |
| pREP3X-cdc17[32-65]-GFP | amino acids 32-65 |

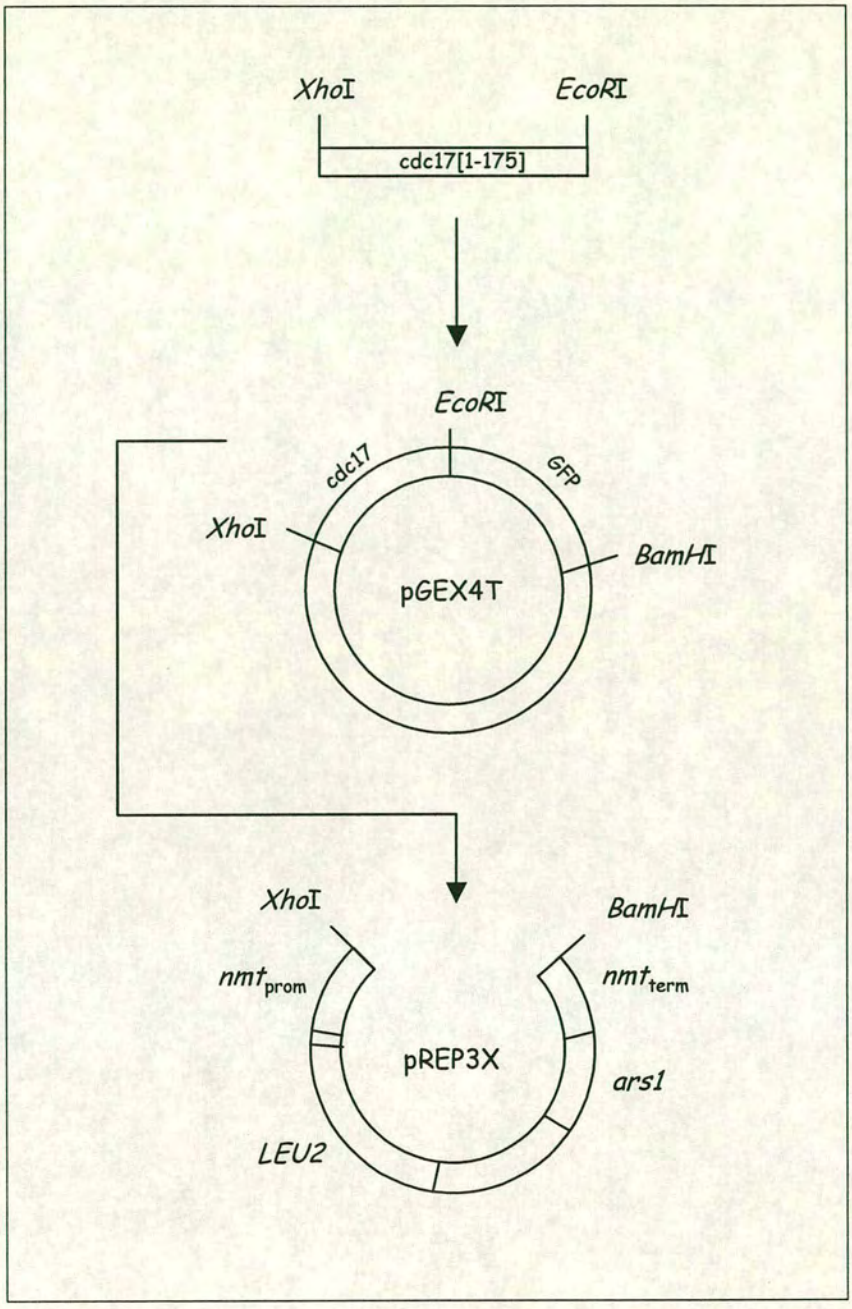



**Figure 6.2**: Cloning strategy for pREP3X-cdc17[1-175]-GFP-constructs

## 6.4 Antibodies and dyes

### 6.4.1 Antibodies

#### 6.4.1.1 Anti-Cdc17 antibodies

Two different antibodies against Cdc17 were raised.

Antisera against a synthetic peptide corresponding to the C-terminal last 20 amino acids of Cdc17 (see Appendix 2), conjugated to KLH/MBS, was produced by Sigma Genosys. Six immunisations and five bleeds were carried out on two rabbits. This antibody was used for Western blotting analyses in this work and will be referred to as anti-Cdc17-peptide antibody. Prior to usage the antiserum was affinity-purified as described in section 6.12.4.

In addition, rabbit antibodies were raised against a GST-fusion of amino acids 391-768 of Cdc17, by Diagnostics Scotland. Testing of the resultant antisera showed that the quality of the anti-Cdc17 antibodies was not as good as the anti-Cdc17-peptide antibodies. Therefore, these antibodies were used only as controls. The generation of the GST-fusion protein, which was used to affinity-purify the above described anti-Cdc17-peptide antibody, is described in section 6.12.4.

#### 6.4.1.2 Secondary antibodies for immunofluorescence

Anti-rabbit Cy3 (code no. 711-166-152) was purchased from Jackson Immunoresearch Laboratories. Alexa488 was a gift from Dr. Hiro Ohkura.

### 6.4.2 Dyes

Mito-Tracker® Red CM-$H_2$Xros was obtained from Molecular Probes (cat.-no. M-7513) and was used to stain *S. pombe* mitochondria.

## 6.5 Bacterial methods

### 6.5.1 Bacterial strains

| strain | genotype | reference |
|---|---|---|
| JM109 | *endA*1, *recA*1, *gyrA*96, *thi*, *hsdR*17 ($r_K^-$, $m_K^+$), *relA*1, *supE*44, Δ(lac-proAB), [F′, *traD*36, *proAB*, *lacI*$^q$*Z* ΔM15] | Yanisch-Perron *et al.*, 1985 |
| CJ236 | F′, cat(pCJ105 ; M13$^s$Cm$^r$)*dut*1, *ung*1, *thi*-1, *relA*1, *spoT*1, *mcrA* | Kunkel *et al.*, 1987 |
| BL21(DE3)pLysS | F$^-$, *ompT*, *hsdS*$_B$, ($r_B^-$, $m_B^-$), *dcm*, *gal*, λ(DE3), pLysS, Cm$^r$ | Studier and Moffat, 1986 |

JM109 was used for propagation of plasmids, CJ236 for oligonucleotide-directed mutagenesis (see section 6.11) and BL21(DE3)pLysS for expression of a GST-fusion protein (see section 6.12.4).

### 6.5.2 Bacterial media

LB (Luria broth): 10 g/l bacto-tryptone
5 g/l bacto-yeast extract
5 g/l NaCl, pH 7.0 (NaOH),
autoclaved and stored at room temperature

For solid media agar was added to a final concentration of 2%.

### 6.5.3 Antibiotics

Where appropriate antibiotics where added to the growth medium as follows:

| antibiotic | stock solution | working concentration |
|---|---|---|
| ampicillin | 100 mg/ml in $H_2O$ | 100 µg/ml |
| chloramphenicol | 34 mg/ml in ethanol | liquid media: 15 µg/ml<br>solid media: 30 µg/ml |
| kanamycin | 10 mg/ml in $H_2O$ | 10 µg/ml |
| tetracycline | 5 mg/ml in ethanol | 5 µg/ml |

### 6.5.4 Transformation

Unless otherwise stated bacteria were routinely transformed with plasmid DNA using the DMSO method. Cells were grown to an OD of 0.5, spun down for 10 minutes at 3000 rpm, resuspended in 1/10 volume of TSB and stored on ice for at least 10 minutes. 100 µl competent cells were added to 0.1-200 ng of DNA and incubated on ice for 30 minutes. 200 µl TSBG were added and the cells were incubated at 37 °C shaking for 1 hour. 100-200 µl were spread on the appropriate media. Transformants containing cloned sequences into pTZ19R were plated out on LB containing ampicillin, X-gal (2% (w/v) stock in DMF, final concentration 0.005%) and IPTG (stock 100 mM, final concentration 100 µM) for blue-white selection.

## 6.6 Fission yeast methods

### 6.6.1 Introduction

Many of the fission yeast methods are based on protocols described in the fission yeast handbook (accessible online: www.bio.uva.nl/pombe/handbook) and Moreno *et al.*, 1991. Further fission yeast methods are described in Alfa *et al.*, 1993, MacNeill and Fantes, 1993 and MacNeill and Fantes, 1997.

### 6.6.2 Strains

The following fission yeast strains were used in this work.

| strain | genotype | reference |
|---|---|---|
| *cdc17-K42* | *cdc17-K42 leu 1-32* | Nasmyth, 1977 |
| *cdc17-L16* | *cdc17-L16 leu 1-32* | Nasmyth, 1977 |
| *cdc17-M75* | *cdc17-M75 leu 1-32* | Nasmyth, 1977 |
| Sp5 | *leu 1-32* $h^-$ | this lab |
| Sp201 | *leu1-32 ura4-D18 his7-366 ade6-M210* $h^+$ | this lab |
| Sp202 | *leu1-32 ura4-D18 his7-366 ade6-M216* $h^-$ | this lab |
| | *cdc17*$^+$*/cdc17*Δ*::kanMX6 leu1-32/leu1-32 ura4-D18/ura4-D18 his7-366/his7-366 ade6-M210/ade6-M216* $h^-/h^+$ | this work |
| | *cdc17*Δ*::kanMX6 leu1-32 his7-366 ade-M21X ura 4-D18*/pUR19-cdc17 | this work |

### 6.6.3 Media and growth conditions

**YE (yeast extract)**

Fission yeast strains were routinely grown in YE. It is a complex medium and inhibits conjugation and sporulation.

5 g/l Oxoid yeast extract

30 g/l glucose

120 mg/l adenine, histidine, leucine, lysine hydrochloride and uracil

**EMM (Edinburgh minimal medium)**

EMM was used as minimal medium for vegetative growth.

3 g/l potassium hydrogen phthallate

2.2 g/l $Na_2HPO_4$

5 g/l NaCl

2% (w/v) glucose

20 ml/l salts (50x stock)

1 ml/l vitamins (1000x stock)

0.1 ml/l minerals (10,000x stock)

salts (50x stock)

52.5 g/l $MgCl_2 \cdot 6H_2O$

0.735 mg/l $CaCl_2 \cdot 2H_2O$

50 g/l KCl

2 g/l $Na_2SO_4$

vitamins (1,000x stock)

1 g/l pantothenic acid

10 g/l nicotinic acid

10 g/l *myo*-inositol

10 mg/l biotin

minerals (10,000x stock)

5 g/l boric acid

4 g/l $MnSO_4$

4 g/l $ZnSO_4 \cdot 7H_2O$

2 g/l $FeCl_2 \cdot 6H_2O$

0.4 g/l molybdic acid

1 g/l KI

0.4 g/l $CuSO_4 \cdot 5H_2O$

10 g/l citric acid

**ME (malt extract)**

ME is a nitrogen limiting medium and plates containing 30 g/l malt extract were used for the induction of conjugation and consequent sporulation.

**Phloxin B**

To check the ploidy of cells, phloxin B was added to solid media. Phloxin B (stock: 5 g/l in $dH_2O$, filter-sterilised) was added to YE after it has cooled below 60°C to a final concentration of 1.75 µg/ml. Phloxin B is a stain that accumulates in dead cells.

Since diploid colonies, because less viable, contain a higher number of dead cells, these colonies will be stained dark pink, whereas haploid colonies appear light pink. Solid media was prepared by adding 2% Difco Bacto Agar.

### 6.6.4 Storage of *S. pombe* strains

For short term storage *S. pombe* strains were kept as patches on solid media at 4°C. For long term storage, glycerol stocks were prepared. Cells are grown up in complete medium to stationery phase. An equal volume of the cell suspension is mixed with 30% glycerol in YE in a cryotube and stored at -70°C.

For re-isolation, material of the glycerol stock is scraped off with a sterile spatula, placed onto a YE plate and incubated at the appropriate temperature. When growth is visible, the material is streaked out to single colonies.

### 6.6.5 Phenotypical analyses

Re-isolated strains need to be checked for their phenotype before carrying out any genetical analyses.

#### 6.6.5.1 Mating type

To determine the mating type of a strain it was crossed (section 6.6.6.1) to $h^+$ and $h^-$ tester strains. After incubation for 3 days the plate was held over iodine crystals. Where sporulation has occurred the cross will be stained black due to the presence of starch in the spores. Also, PCR using mating type specific primers can be used (see section 6.9.3).

#### 6.6.5.2 Temperature sensitivity

Temperature-sensitivity was tested by replica-plating patches of the re-isolated strain onto YE and incubating at the permissive and restrictive temperature. After

incubation for several days, the plates are examined for growth. The permissive temperature for used *cdc17*$^{ts}$ strains is 25°C and the restrictive temperature is 36°C.

#### 6.6.5.3 Auxotrophy

To test for auxotrophy patches of the re-isolated strain on YE were replica-plated onto minimal medium with and without the appropriate supplement. After incubation for several days, the plates are examined for growth. The auxotrophic markers used in this study are leucine, adenine, histidine and uracil. Where appropriate, stock solutions (adenine, leucine: 7.5 g/l and histidine, uracil 3.75 g/l) were added to media to a final concentration of 112.5 mg/l.

### 6.6.6 Genetic analyses

#### 6.6.6.1 Genetic crosses

Strains were crossed by mixing together freshly isolated cells of opposite mating type on a ME plate. A loopful of sterile $H_2O$ is then used to thoroughly mix the cells. The plate is incubated at 25°C for several days until zygotic asci are visible under the microscope and/or the mating mix turns blue black upon exposure to iodine vapour (see section 6.6.5.1). The progeny of the crosses are examined by tetrad analysis or random spore analysis.

#### 6.6.6.2 Diploid construction

In this work, the isolation of diploids was aided by the use of two complementing adenine mutants. Strains bearing the *ade6-M210* or *ade6-M216* alleles are adenine auxotroph and colonies appear pink due to the accumulation of a metabolic precursor. A heterozygous diploid carrying both alleles is prototrophic for adenine and colonies appear white.

#### 6.6.6.3 Random spore analysis

A loopful of 2-3-day-old mating mix was resuspended in 500 µl sterile $H_2O$, 10 µl of a 1:10 dilution of helicase (*Helix pomatia* juice) is added and the mixture is incubated overnight at 28°C. The mix was checked microscopically for ascus wall breakdown and $10^{-1}$ – $10^{-4}$ dilutions were plated out on YE. The arising colonies were then analysed further.

#### 6.6.6.4 Tetrad analysis

A loopful of a 2-day-old cross was streaked out in a line at one side of a YE plate. The asci of the cross were pulled using a micromanipulator (Singer Instruments, UK). The plate was incubated for several hours at 32°C or overnight at 16°C, until the asci walls have broken down. Each ascus was then dissected and the four spores were separated and placed into a line using the micromanipulator. The spores are incubated at the appropriate temperature until colonies form.

### 6.6.7 Physiology

#### 6.6.7.1 Growth of liquid cultures

Exponentially growing cells were obtained by inoculating a freshly growing, single colony into 10 ml of appropriate medium and incubating for 2-4 days until the cells are in stationery phase. This preculture was then diluted into fresh media so that overnight incubation with agitation yields a culture with an OD of 0.2 – 0.4. (an OD of 0.05 corresponds to ~$10^6$ cells/ml.) Exponentially growing cells usually have an OD of between 0.1 and 0.5. Incubation temperature is generally 32°C, but can vary between 25°C and 36°C.

#### 6.6.7.2 Measurement of cell numbers

Cell numbers were determined using a Beckman Coulter Counter Z2. 100 µl cell suspension was added to 10 ml of Isoton and mixed. The sample was sonicated for

10 seconds at 5% with a high intensity ultrasonic processor (Sonics & Materials), mixed again and counted with the Coulter Counter according to the manufacturer's instructions. The average of two counts was used for calculation of cell number.

#### 6.6.7.3 Measurement of cell lengths

1 ml of exponentially growing cells were spun down for 10 seconds and the supernatant was discarded. 2 µl of cell suspension was pipetted onto a glass slide, left to dry briefly before covering. Cells were viewed at 40x magnification. For cell length measurement an eyepiece graticule was used, where the smallest unit corresponds to 2.5 µm. Cell length was measured at division, i.e. those cells, that had clearly formed a septum, but had not yet undergone cytokinesis, were considered. 25-35 cells were examined and average cell length and standard deviation were calculated using Microsoft Excel.

### 6.6.8 Preparation of *S. pombe* chromosomal DNA

Fission yeast chromosomal DNA was prepared essentially as described in the fission yeast handbook (www.bio.uva.nl/pombe/handbook).

10 ml of the appropriate media was inoculated with a single yeast colony and incubated for 1-2 days at the required temperature until the culture reaches stationary phase. Cells were harvested and resuspended in CPS buffer containing 2.5 mg/ml Zymolase 20-T (ICN). The cells were transferred to a 1.8 ml microfuge tube and incubated at 37°C for 60 minutes. Cells were harvested and resuspended in 300 µl 5x TE. 35 µl 10% SDS (w/v) was added and the suspension was incubated at 65°C for 5 minutes. 100 µl 5M potassium acetate was added and incubated on ice for 30 minutes. After centrifugation for 15 minutes at 4°C the supernatant (~400 µl) was added to 1 ml ice-cold ethanol and mixed. It was centrifuged for 10 minutes at 4°C and the pellet was resuspended in 400 µl of 5x TE containing 100 µg/ml bovine pancreatic RNAse A and incubated at 37°C for 2-4 h. The DNA was extracted once with 400 µl phenol, once with 400 µl phenol-chloroform-isoamylalcohol (25:24:1) and once with 400 µl chloroform. Approximately 400 µl of the aqueous layer was

transferred into a new tube and 40 µl 5 M potassium acetate, pH 5.6, and 1ml ice-cold ethanol was added and mixed. The DNA was precipitated with a centrifugation for 30 minutes at 4°C and it was washed once with 200 µl 70% ethanol. The pellet was dried under vacuum and routinely resuspended in 100 µl of TE, pH 8.0.

### 6.6.9 *S. pombe* transformation

Routinely *S. pombe* was transformed by electroporation based on the method described by Prentice (1992).

200 ml cultures of *S. pombe* strains to be transformed were grown to late-logarithmic phase (OD 0.4), washed three times in ice-cold sorbitol and resuspended in ice-cold sorbitol at 1 x $10^9$ cells/ml. 200 µl of cell suspension was mixed with 0.5-1 µg of plasmid DNA and immediately transferred to an ice-cold electroporation cuvette (0.2 cm electrode gap). Cells were pulsed at 2.5 kV (11.25 kV/cm), 200Ω and 25 µF in a Bio-Rad Gene Pulser™ II. Immediately after the pulse 0.5 ml of ice-cold sorbitol was added to the cuvette. 200 µl were plated on selective media and incubated at 32°C for five days. Usually >1000 transformants are obtained with this protocol.

### 6.6.10 Gene replacement

Deletion of chromosomal cdc17 was carried out according to the method of Bähler *et al.*, 1998. In a diploid *S. pombe* strain, one *cdc17* allele was replaced with *kanMX6*, the gene product of which confers resistance to the antibiotic G418 (Gibco-BRL).

The *kanMX6* module was amplified from plasmid pFA6a-kanMX6 by standard DeepVent-PCR with primers that terminate in short stretches of homology (79 bp) to the genomic sequences immediately upstream and downstream of the *cdc17* coding sequence (primers #3 and #4, respectively, see section 6.2). 15 µg of PCR product was ethanol-precipitated, resuspended in 10 µl TE and transformed into the diploid strain of genotype *leu1-32/leu1-32 ura4-D18/ura4-D18 his7-366/his7-366 ade6-M210/ade6-M216* $h^-/h^+$, which was obtained by crossing the strains Sp201 and

Sp202 (see section 6.6.2) and selecting for $ade^+$ diploids. The transformation protocol was based on the lithium acetate procedure (Keeney and Boeke, 1994). For this, a 20-ml-culture was grown to mid-log phase (~$10^7$ cells/ml) at 32°C. Cells were washed once in 20 ml $H_2O$, the pellet was resuspended in 1ml $H_2O$, transferred to an Eppendorf tube and washed once with 100 mM LiAc/TE (pH 7.5). The cell pellet was resuspended in LiAc at 2 x $10^9$ cells/ml. 100 µl cells were mixed with 2 µl salmon sperm carrier DNA (10 mg/ml, Clontech) and 15 µg of transforming DNA and incubated at room temperature for 10 min. 260 µl of 40% PEG/LiAc/TE (PEG 4000) was added, the cell suspension mixed gently and incubated at 30°C for 30-60 min. 43 µl of DMSO were added and the cells were heat-shocked at 42°C for 5 min. Cells were washed with 1ml $H_2O$, resuspended in 0.5 ml $H_2O$ and 2 x 250 µl were plated onto YE. The plates were incubated for ~18 h at 32°C resulting in a lawn of cells and then replica-plated on YE containing 100 mg/l G418. The plates were incubated for 3 days at 32°C and appearing big colonies were re-streaked onto YE + G418.

Verification of correct integration events was carried out by standard Taq-PCR. Genomic DNA from four stable transformants and the corresponding wild-type strain was prepared (section 6.6.9) and used as a template in a PCR reaction with different combinations of primers. Two primers corresponded to sequences within the transforming fragment going in either direction (primers #6 and #8) and two primers (primers #5 and #7) covered a region upstream and downstream of the *cdc17* coding sequence to be replaced (see also Chapter 2, primers see section 6.2).

## 6.6.11 Plasmid shuffle

The basis for the plasmid-shuffle experiments would be a haploid *cdc17*Δ which is maintained by a wild-type copy of the gene on a plasmid carrying the $ura4^+$ gene complementing the uracil auxotrophy of the strain. After transformation of pREP41X-cdc17-constructs ($leu^+$) into *cdc17*Δ/pUR19-cdc17 ($ura^+$) cells carrying both plasmids can be plated out on 5-FOA (5-fluoroorotic acid). 5-FOA selects against $ura^+$ cells (Boeke *et al.*, 1984, Grimm *et al.*, 1988) and therefore the ability of

truncated or mutated *cdc17* versions, expressed from pREP41X, to rescue the deletion can be assessed. Solid 5-FOA was added to agar media, held at 55°C, and dissolved for 10 minutes at 55°C (final concentration 1 mg/ml).

The spores from the heterozygous deletion diploid *cdc17*$^{+}$/*cdc17*Δ::*kanMX6 leu1-32/leu1-32 ura4-D18/ura4-D18 his7-366/his7-366 ade6-M210/ade6-M216 h*$^{-}$/*h*$^{+}$ that had been plated out on EMM + supplements lacking leucine in the presence or absence of thiamine were tested for their phenotype. An isolate which was G418$^{res}$, leu$^{+}$, ade$^{-}$ (i.e. haploid) was transformed with pUR19-cdc17 and it was selected for ura$^{+}$ leu$^{-}$ colonies. The resulting isolate *cdc17*Δ::*kanMX6 leu1-32 his7-366 ade-M21? ura4-D18*/pUR19-cdc17 was used for complementation experiments.

After transforming in pREP41X-cdc17-constructs, cells were plated out on media selecting for both plasmids and either containing or lacking thiamine, corresponding to subsequent culturing and plating steps. Well-isolated transformants were replicated at least twice on uracil containing medium in order to dilute out the *ura4*$^{+}$ gene product before growing up in liquid culture and plating out 1000 cells on media containing 5-FOA and uracil, either with or without thiamine (as before) and supplemented with histidine and adenine.

## 6.7 Fluorescence microscopy

### 6.7.1 GFP-fusion proteins

GFP is a protein from the jellyfish *Aequorea victoria* which fluoresces green (green fluorescent protein) and it is widely used as a marker for gene expression *in vivo* and as a cytological marker fused to protein-coding sequences (Sawin, 1999). For example, GFP-fusion genomic DNA libraries were constructed and screened for intracellular protein localisation in *S. pombe* as a means of identifying the function of unknown proteins (Sawin and Nurse, 1996, Ding *et al.*, 2000).

In this work, a mutant form of GFP, namely GFP[S65T], was used, which results in a six- to ten-fold enhanced fluorescence compared to the wild-type GFP protein (Heim *et al.*, 1995, Sawin, 1999). For convenience, GFP[S65T] will be referred to simply as GFP throughout this work.

Expression of GFP-fusion proteins were undertaken in *S. pombe leu1-32 h⁻* cells (strain Sp5). Precultures of cells carrying the respective expression vectors (see section 6.3.3.2) contained thiamine in order to avoid toxic effects due to overproduction. After inoculating a main culture, cells were grown to an OD of 0.1-0.2. For viewing under the microscope, 1 ml of cell suspension was spun down briefly and ~2 µl of the loose pellet was pipetted onto a glass slide, dried and covered with a cover slip.

When staining with Mito-Tracker® was carried out, 1 ml cell suspension was spun down, resuspended in 200 nM Mito-Tracker® (in EMM) and incubated in the dark for 30 minutes. Cells were washed once in EMM and prepared for microscopy as above. Cells were viewed, using green and red filter, under a Zeiss Axioplan 2 microscope with a 100x oil-immersion objective. Photos were taken using the OpenLab software and images were processed in AdobePhotoshop.

## 6.7.2 Immunofluorescence

*S. pombe* Sp5 cells transformed with pREP41X vectors encoding several versions of *cdc17* and grown in the absence of thiamine, were subjected to immunofluorescence. Generally, this method involves fixation of cells, cell wall digestion, primary and secondary antibody application and microscopy. Cells were viewed under a Zeiss Axioskop microscope with a 66x oil-immersion objective. Pictures were taken using the Smart Capture VP software and processed in Adobe Photoshop.

Whenever cells were spun down in microfuge tubes in the following protocol, the double spin technique was used in order to minimise cell loss: cells were spun down for one minute, then the orientation of the tube was reversed, and cells were spun down again for one minute.

**Solutions**

PEM: 100 mM Pipes
1 mM EGTA
1 mM $MgSO_4$, pH 6.9

PEMBAL: PEM
+ 1% BSA
+ 0.1% sodium azide
+ 100 mM L-lysine HCl

PEMS: PEM + 1.2 M sorbitol

Formaldehyde solution (for fixation, prepare freshly, TOXIC):

For 10 ml of a 30% solution, 3 g para-formaldehyde was placed in a 50 ml tube and 10 ml 65°C-prewarmed PEM was added. The solution was incubated at 65°C for 5 minutes and 12 drops of 5 M NaOH were added. The solution was mixed, incubated at 65°C until all of the formaldehyde has gone into solution, cooled on ice to 32°C and spun down briefly before use.

**Fixation**

2.5 ml formaldehyde solution was placed in a 50 ml tube and 20 ml cell culture (OD ~0.2) was added quickly and mixed thoroughly. 0.5-1 minute later glutaraldehyde was added to a final concentration of 0.2%, the tube was mixed well and incubated on a shaker for 90 minutes. Fixation was stopped by spinning down the cells. The

cells were resuspended in 1 ml PEM, transferred to a microfuge tube and washed three times in PEM.

**Cell wall digestion**

Cells were resuspended in 1 ml PEMS + 2.5 mg Zymolase-20T and incubated at 37°C until ~90% of cells have become spheroplasts. Cells are pelleted and resuspended in 1% Triton-X100 in PEMS. After 0.5 minute cells are washed three times in PEM.

**Quenching**

This procedure is aimed at reducing free aldehyde groups available to interact with the incoming antibodies. A little sodium borohydride was placed in a 50 ml tube, 5 ml PEM was added and mixed (solution will bubble extensively). 500 µl sodium borohydride solution was added to the cell pellet and incubated at room temperature for five minutes with the caps left open. Cells were pelleted and washed twice in PEM, resuspended in PEMBAL and incubated on a rotating wheel at room temperature for 30 minutes.

**Staining**

An aliquot of cells was pelleted, resuspended in 100 µl of a primary antibody solution in PEMBAL at the desired dilution and incubated overnight on a rotating wheel. The cells were pelleted, washed three times in PEMBAL, resuspended in 100 µl of secondary antibody solution at the desired dilution and incubated five hours on a rotating wheel.

**Microscopy**

Cells were pelleted, washed once in PEMBAL, once in PBS and resuspended in PBS containing 0.2 µg/ml DAPI (stock: 20 µg/ml). Cells were pelleted and resuspended in PBS + 0.1% sodium azide. The cells are mounted on poly-lysine coated coverslips and left to air-dry. The coverslips are placed onto a glass slide with 5 µl mounting medium (10 µl of 10 mg/ml paraphenylene diamine in PBS + 90 µl glycerol).

## 6.8 General DNA methods

### 6.8.1 Phenol-chloroform extraction

DNA was extracted from proteinaceous solutions with phenol/chloroform.
An equal amount of phenol was added to the DNA solution, vortexed for one minute and spun down at max. speed in a microfuge for 5 minutes. The upper aqueous layer was transferred to a tube containing an equal amount of phenol:chloroform:isoamylalcohol (25:24:1) and treated as above. The upper aqueous layer was then added to a fresh tube with the same volume of chloroform. The aqueous layer was transferred in a fresh tube and the DNA was recovered by ethanol precipitation.

### 6.8.2 Ethanol precipitation

DNA solutions were concentrated and/or purified by precipitation. 2.5 volumes of ethanol (-20°C) and 0.1 volume of 3M sodium acetate, pH 5.2, was added to the DNA, mixed thoroughly and precipitated at –20°C for 30 minutes. The DNA was pelleted by centrifugation at max. speed for 30 minutes at 4°C in a microfuge. The DNA was then washed in 70% ethanol, dried under vacuum and resuspended in a suitable amount of TE, pH 8.0, or $H_2O$.

### 6.8.3 Preparation of plasmid DNA

Small scale plasmid preps were routinely prepared with the Promega Wizard plasmid prep kit (cat.-no. #A1460). Typically, 100 µl at 100 ng/µl were obtained. Plasmid DNA which is to be subsequently sequenced is prepared using the Promega syringe barrel columns (cat.-no. #A7510) followed by an ethanol precipitation. The DNA is typically resuspended in 30 µl sterile $dH_2O$ and the yield usually is ~50 ng/µl. The QIAGEN filterMidi (cat.-no. 12243) or filterMaxi (cat.-no. 12262) kits were

routinely used for large scale plasmid preps, yielding 100 µl at between 1-5 µg/µl. In all cases, plasmid DNA was prepared according to the manufacturer's instructions.

### 6.8.4 Restriction enzyme digests of DNA

Restriction endonuclease digestion of DNA was carried out in 20-µl-reactions for analytical digests and 100-µl-digests for preparative digests. The restriction enzymes were used as recommended by the manufacturers with a digest typically containing between 200 ng and 3 µg of DNA and between 2 and 5 units of restriction enzyme in the supplied 1x reaction buffer. The digest was incubated at the recommended temperature, normally for a period of 2 h. The products of the digestion were either analysed by agarose gel electrophoresis (section 6.8.7) or purified using a QIAquick® gel extraction column (cat.-no. 28704).

### 6.8.5 Removal of phosphates from DNA ends

The removal of 5' phosphate groups of linearised plasmid DNA was carried out to avoid re-circularisation of the plasmid vector DNA in subsequent ligation reactions. Typically, 0.5 µl of Shrimp Alkaline Phosphatase (USB, 1 unit/µl) was added to 2-3µg digested plasmid DNA and incubated for 30 minutes at 37°C. The digested and phosphatased vector was purified as described in the previous section.

### 6.8.6 Ligations

Ligations were routinely carried out in a final volume of 10 µl, containing 0.5-1.0 µg total DNA with T4 DNA ligase (Roche) in the supplied reaction buffer. Vector and insert DNA were present in an approximately 1:3 molar ratio. Ligation reactions were incubated overnight at 16°C and stopped by freezing at -20°C.

### 6.8.7 Agarose gel electrophoresis

DNA fragments produced by restriction enzyme digest or by PCR amplification was analysed in 0.8-1.5% (w/v) agarose gels. Gels were prepared by melting the appropriate amount of agarose in 1x TAE and adding ethidium bromide to a final concentration of 0.1 µg/ml. The gel running buffer consisted of 1x TAE with 0.02 µg/ml ethidium bromide. Gels were run at a voltage of 4-8 V/cm for 1.5 h. DNA samples were mixed with 6x loading buffer.

The size of the DNA fragments and its quantity was compared to molecular weight markers. Typically, 10 µl of Gibco-BRL 1 kb DNA ladder (cat.-no. 15615-016, prepared according to the manufacturer's instructions) was loaded alongside the samples (1.6 kb band contains 100 ng DNA).

The DNA was visualised using a UV-transilluminator and pictures were taken with a UVP DOC-IT™ camera (Ultra-violet Products).

### 6.8.8 Purification of DNA from agarose gel slices.

DNA was purified from agarose gels using the QIAquick® Gel Extraction Kit (Qiagen). DNA fragments were separated by agarose gel electrophoresis and the bands visualised on a UV transilluminator (set to low intensity). The band of interest was excised with a clean scalpel and purified according to the manufacturers' protocol. DNA was typically eluted in 30µl of sterile, distilled water.

## 6.9 DNA amplification by PCR

### 6.9.1 General

Polymerase chain reaction (PCR) was routinely used to amplify DNA fragments from template DNA, which either was plasmid DNA or *S. pombe* genomic DNA. Routinely, reaction volumes were 20 µl for analytical and 100 µl for preparative PCR reactions. If the PCR products were to be analysed for the presence of a specific

DNA fragment the robust Taq DNA polymerase (Roche, cat.-no. 1146165) was used. In case the amplified DNA fragment was subsequently to be cloned into a vector the Deep Vent DNA Polymerase (NEB, cat.-no. #M0258), which possesses exonuclease activity and therefore minimizes errors during the amplification procedure, was used. If Deep Vent was the polymerase of choice the amount of additional $Mg_2SO_4$ to add to the reaction mixture had to be optimized for each DNA template. 0.5 units Taq Polymerase was used and DeepVent Polymerase was added at 0.4 units.
A typical 20 µl reaction was set up as follows:

| | |
|---|---|
| DNA | 20-200 ng |
| 5' oligonucleotide primer [10µM] | 1 µl |
| 3' oligonucleotide primer [10µM] | 1 µl |
| 10 x polymerase buffer | 2 µl |
| dNTPs [10 mM] | 0.4 µl |
| $Mg_2SO_4$ (if using DeepVent) | 0-6 mM |
| DNA Polymerase: Deep Vent | 0.2 µl |
| or Taq | 0.1 µl |
| dH20 | ad 20 µl |

0.2 ml tubes were used for the PCR reaction mix and the reactions were carried out in a "Genius" PCR machine from Techne. The standard PCR programme was as follows:

| | | |
|---|---|---|
| Step 1 Denaturation: | 94 °C for 1.5 minutes | |
| Step 2 Primer annealing: | 55 °C for 1.5 minutes | |
| Step 3 Extension: | 72 °C for 3.0 minutes, | 30 cycles |

## 6.9.2 Colony PCR

If *E. coli* or *S.pombe* colonies are to be screened directly for the presence of a certain DNA fragment a crude DNA preparation as template for a PCR reaction can be made

as follows: The equivalent of a 1-2 mm colony was denatured in 10 µl $H_2O$ or 0.02 N NaOH at 95 °C for 5 minutes. 2 µl were used in a PCR reaction.

### 6.9.3 Mating type PCR

The mating type of a fission yeast strain can be determined using mat1-P and mat1-M specific primers. Common primer #MT1, the mat1-P specific primer #MP and the mat1-M specific primer #MM produce characteristic PCR products of 987 bp and 729 bp for $h^+$ and $h^-$ cells, respectively. The sequence of the mating type primers can be found in section 6.2 and the conditions for the PCR reaction are detailed below.

| PCR reaction mixture | | PCR cycle | |
|---|---|---|---|
| DNA | 2.0 µl | Step 1 | 94 °C for 4.0 minutes |
| #MT1 [10µM] | 2.0 µl | Step 2 | 94 °C for 0.5 minutes |
| #MP [10µM] | 2.0 µl | Step 3 | 52 °C for 0.5 minutes |
| #MM [10µM] | 2.0 µl | Step 4 | 72 °C for 2.0 minutes |
| 10 x polymerase buffer | 2.0 µl | | step 2 - 4 for 30 cycles |
| dNTPs [10 mM] | 0.4 µl | Step 5 | 72 °C for 5.0 minutes |
| H2O | 9.4 µl | | |
| Taq polymerase | 0.2 µl | | |
| Total | 20 µl | | |

## 6.10 DNA sequencing

### 6.10.1 General

Plasmid DNA to be sequenced was prepared using the Promega syringe barrel columns and quantitated by visualisation on an agarose gel. Reactions were

performed with the dRhodamine or bigDye terminator cycle sequencing kit (Perkin Elmer). A typical reaction mix was set up as follows:

| | |
|---|---|
| Template DNA (100-250 ng) | ≤ 4.0µl |
| Terminator mix | 4.0µl |
| Primer (1.6 pmol) | 1.0µl |
| $H_2O$ | ad 10 µl |

Step 1: 96°C for 30 seconds
Step 2: 50°C for 15 seconds
Step 3: 60°C for 4 minutes, 25 cycles

DNA was precipitated from the reaction mix by adding 25µl of ethanol and 1µl of 3M sodium acetate, pH 5.2, incubating on ice for 15 minutes followed by centrifugation at 14,000 rpm for 30 minutes at 4°C. The pellet was washed with 125µl of 70% (v/v) ethanol and dried under vacuum. Samples were run by the ICMB Sequencing Service (University of Edinburgh) on an ABI PRISM 377 DNA sequencer and the sequence analysed using the Gene Jockey II programme on a Macintosh computer.

### 6.10.2 Sequencing of *cdc17*$^{ts}$ alleles

Genomic DNA was prepared from the temperature-sensitive strains *cdc17*-K42, *cdc17*-L16 and *cdc17*-M75 strains. The *cdc17* gene was amplified from the genomic DNA by DeepVent-PCR using primer #1 and primer #2, which results in a DNA fragment carrying *cdc17* with *BamH*I sites at either end. The PCR products were digested with *BamH*I and cloned into pTZ19R. The resultant plasmids were used as template for DNA sequencing with plasmid-specific primers #pTZH and #pTZE and *cdc17*-specific primers #OSeq, #1Seq, #2Seq, #3Seq, #4Seq and #5Seq.

## 6.11 Site-directed mutagenesis

### 6.11.1 *In vitro* oligonucleotide-directed mutagenesis

Most truncations and mutations were engineered using the MutaGene Phagemid Kit version 2 (Bio-Rad 170-3581).

pTZ19R-cdc17 was transformed into the *dut ung* strain CJ236. The *dut* mutation inactivates dUTPase and cells accumulate high levels of dUTP which is incorporated into nascent DNA. The *ung* mutation inactivates uracil N-glycosylase, which allows uracil to remain in the DNA. Phagemid particles, whose single-stranded DNA contains some uracils are produced from these cells by superinfection with the helper phage M13K07. This DNA is purified and used as the template in an *in vitro* mutagenesis reaction primed by an oligonucleotide containing the desired mutation. The resulting double-stranded DNA is transformed into JM109 which contains a functional uracil N-glycosylase. The uracil-containing strand is strongly selected against and the non-uracil containing mutagenic survivor is left to replicate. Normally, mutation efficiencies were ≥ 50%.

Truncations of Cdc17 were constructed by performing the *in vitro* mutagenesis reactions with oligonucleotides introducing *BamH*I recognition sequences at the appropriate sites within *cdc17*. After the mutations were confirmed by sequencing, *BamH*I restriction digests were performed and the desired fragment was subcloned from pTZ19R into pREPX vectors. Mutageneses reactions were performed essentially following the manufacturer's instructions. The oligonucleotides used to generate the different constructs are listed below.

### 6.11.2 Mutagenesis by PCR

PCR reactions generating mutagenic *cdc17* sequences or introducing restriction sites were carried out as described in section 6.9.1. The PCR products were subsequently run on an agarose gel, excised from the gel and cleaned up. After restriction digest of the product fragment, it was cloned into the target vector using standard procedures.

## 6.11.3 Oligonucleotides used for mutageneses

The following oligonucleotides were used to create truncations or mutations in the *cdc17* sequence. Oligonucleotide sequences corresponding to the *cdc17* sequence are indicated as numbers (see Appendix 1) and mutagenic sequences are written out. Restriction sites are underlined. Since the initial preparation of single-stranded phagemid DNA was generated from pTZ19R-cdc17, where the insert was in the wrong orientation, primers used for *in vitro* mutagenesis are in the reverse direction. Constructs which are not mentioned below were engineered through subcloning of appropriate restriction digest fragments.

| construct | method | tmpl* | sequence |
|---|---|---|---|
| cdc17-NΔ19 | ivM | 1) | 1298-1280 AACAAGGATCC 1268-1247 |
| cdc17-NΔ65 | ivM | 1) | 1478-1458 CATAACAAGGATCC 1443-1425 |
| cdc17-NΔ96 | ivM | 1) | 1573-1551 CATAACAAGGATCC 1536-1515 |
| cdc17-NΔ125 | ivM | 1) | 1657-1638 CATAACAAGGATCC 1623-1603 |
| cdc17-NΔ390 | ivM | 1) | 2516-2495 AACAAGGATCC 2483-2460 |
| cdc17-K416A | ivM | 1) | 2587-2572 GC 2569-2559 |
| cdc17-M20A | ivM | 1) | 1296-1282 GC 1279-1268 |
| cdc17-Δ31[1-10rfc1] | ivM | 1) | 1374-1356 AAAGAAGCTTCGAATGTCAGAA<br>TTACTCATAACAAGGATCCCCGGGTACCG |
| cdc17-NΔ31 | PCR | 2) | GATCGGATCCTTGTT 1356-1379, F<br>GATCGGATCC 3636-3608, R |
| cdc17[32-175]-GFP | PCR | 3) | GGCCCTCGAGATTGTTATG 1356-1379, F<br>GGCCGAATTC 1787-1763, R |
| cdc17[20-175]-GFP | PCR | 4) | GGCCCTCGAGATACGG 1280-1300, F<br>GGCCGAATTC 1787-1763, R |
| cdc17[1-175]-M20A-GFP | PCR | 5) | GGCCACTCGCCTCGAGTTGTT 1223-1241, F<br>GGCCGAATTC 1787-1763, R |
| cdc17[66-175]-GFP | PCR | 4) | GGCCACTCGCCTCGAGATTGTTATG 1458-1478, F<br>GGCCGAATTC 1787-1763, R |
| cdc17[97-175]-GFP | PCR | 4) | GCCACTCGCCTCGAGATTGTTATG 1551-1573, F<br>GGCCGAATTC 1787-1763, R |
| cdc17[125-175]-GFP | PCR | 4) | GCCACTCGCCTCGAGATTGTTATG 1638-1659, F<br>GGCCGAATTC 1787-1763, R |
| cdc17[32-65]-GFP | PCR | 3) | GGCCCTCGAGATTGTTATG 1356-1379, F<br>CCGGGTCGAC 1457-1429, R |

| | | | |
|---|---|---|---|
| cdc17 [1-175] ΔNLS-GFP | PCR | 3) | initially two PCR reactions, which overlap at the mutation site, where a NotI restriction site is introduced<br>upper: GGCCACTCGCCTCGAGTTGTT 1223-1241, F<br>GCACCATTACGCGGCCGC 1445-1421, R<br>lower: GCGATTACCAGCGGCCGC 1454-1479, F<br>GGCCGAATTC 1787-1763, R |
| cdc17 [1-175] M20AΔNLS-GFP | PCR | 5) | two PCR reactions, which overlap at the mutation site, where a NotI restriction site is introduced<br>upper: GGCCACTCGCCTCGAGTTGTT 1223-1241, F<br>GCACCATTACGCGGCCGC 1445-1421, R<br>lower: GCGATTACCAGCGGCCGC 1454-1479, F<br>GGCCGAATTC 1787-1763, R |

ivM = *in vitro* mutagenesis

* templates:

1) pTZ19R-cdc17 was used to produce single-stranded phagemid DNA (ivM) for subsequent *in vitro* oligonucleotide-directed mutagenesis.

For PCR reactions the following plasmids were used as templates:

2) pTZ19R-cdc17

3) pREP3X-cdc17[1-175]-GFP

4) pGEX4T-cdc17[1-175]-GFP

5) pREP41X-cdc17-M20A

## 6.12 Protein methods

### 6.12.1 *S. pombe* small scale protein extracts

*S. pombe* denatured protein extracts were routinely prepared using the Ribolyser™ (Hybaid, cat.-no. 6000220/110).

1. *S. pombe* cultures were grown to an $OD_{600}$ of 0.2 - 0.4.
2. Samples with equal OD (usually 10 ml at $OD_{600}$ = 0.3) were taken and cells were pelleted at 3000 rpm for 4 minutes.
3. Cells were resuspended in 5 ml ice-cold STOP buffer (see below) and spun down at 3000 rpm for 4 minutes.
4. Cells (well drained) were resuspended in 100 µl buffer A (see below).
5. Cell suspension was added to Ribolyser tube and vortexed 2 seconds to mix.
6. Tubes were fitted in the Ribolyser and processed at 4.0 g for 20 seconds.
7. 100 µl of 2x SDS sample buffer were added to the bottom of the tube, vortexed 2 seconds to mix and heated to 95 °C for 5 minutes.
8. After tubes had cooled down their base was punctured, clipped into a 2 ml microfuge tube, placed into a 50 ml Falcon tube and centrifuged at 3000 rpm for one minute
9. Prior to gel loading the protein sample was spun down for 3 minutes at maximum speed.

| STOP-buffer | buffer A |
|---|---|
| 150 mM NaCl | 10 mM sodium phosphate buffer, pH 7.0 |
| 50 mM NaF | 100 µl Complete Inhibitors (10x) |
| 10 mM EDTA | (Roche, cat.-no. 1 836 1563) |
| 1 mM $NaN_3$ | 1% Triton X-100 |
| | 0.1% SDS |
| | 1mM EDTA |
| | 150 mM NaCl |
| | 1mM PMSF (add immediately before use) |
| | $H_2O$ ad 1 ml |

## 6.12.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Discontinuous SDS polyacrylamide gels based on the method described by Laemmli (1970) were used to separate proteins electrophoretically. Gels were poured using the Mini-PROTEAN II slab cell for miniature polyacrylamide gels (BioRad 165-2940). Discontinuous polyacrylamide gels consist of an upper stacking gel, where the proteins become concentrated, which results in a better resolution in the lower separating gel. The separating gels had acrylamide concentrations of between 7% and 12% depending on the molecular weights of the proteins to be separated. Below the recipes for 7% and 12% separating gels and the stacking gel are given, which is enough to pour two mini-gels. The acrylamide solution contained 30% (w/v) acrylamide mixed with bis acrylamide at a ratio of 37.5:1 (Severn Biotech Ltd).

| | separating gel, 7% | separating gel, 12% | stacking gel, 5% |
|---|---|---|---|
| 1.5 M Tris-HCl, pH 8.8 | 2.5 ml | 2.5 ml | --- |
| 0.5 M Tris-HCl, pH 6.8 | --- | --- | 2.5 ml |
| acrylamide solution | 2.5 ml | 4 ml | 1.33 ml |
| 10% (w/v) SDS | 100 µl | 100 µl | 100 µl |
| $H_2O$ | 4.85 ml | 3.35 ml | 6.1 ml |
| 10% (w/v) AMPS | 50 µl | 50 µl | 50 µl |
| TEMED | 5µl | 5 µl | 10 µl |
| | 10 ml | 10 ml | 10 ml |

The gel apparatus was assembled according to the manufacturers instructions. The separating gel solution was mixed and poured. To assure the formation of an even surface the separating gel is overlayed with some drops of $H_2O$. The gel was allowed to polymerise for ca. 45 minutes. After removal of $H_2O$ the prepared stacking gel solution was poured and the combs fitted. After 45 minutes the combs were removed and the wells were washed with protein gel running buffer. The gels were fitted in the buffer tank and the tank was filled with protein gel running buffer.

Protein samples were mixed with an equal volume of 2x SDS sample buffer, boiled for 5 minutes, cooled down and briefly spun down before loading on the gel.
As a molecular marker broad range pre-stained protein standards (NEB, cat.-no. 7708S ,size range 6.5–175 kDa) was used.

### 6.12.3 Western blotting

The Mini Trans blot cell (BioRad, cat.-no. 170-3930) was used for electrophoretically transferring proteins from SDS polyacrylamide gels to PVDF membranes (BioRad, 162-0184). The pads, blotting paper (Whatman 3MM), membranes and gels were soaked in chilled Transfer buffer for 15 minutes. The blotting apparatus was assembled according to the manufacturers' instructions and fitted into the tank. A cooling block was added and the tank was filled with transfer buffer, which was stirred throughout the transfer, The transfer was allowed to proceed for one hour at 100 V. The proteins immobilized on the membranes were detected using the ECL Chemiluminescence system (Amersham Pharmacia, RPN 2209). After electroblotting the membranes were treated as follows:

1. The membranes were incubated in 5% (w/v) dried milk in PBST for one hour with agitation
2. The membranes were briefly rinsed twice with PBST and washed in large volumes (approx. 50 ml) of PBST: 1 x 15 minutes, 3 x 5 minutes.
3. The membranes were transferred in small perspex trays and incubated in 10 ml primary antibody solution (antibody diluted in PBST). Incubation was allowed to proceed overnight at 4 °C with gentle shaking.
4. The membranes were briefly rinsed twice with PBST and washed in large volumes (approx. 50 ml) of PBST: 1 x 15 minutes, 3 x 5 minutes.
5. The membranes were transferred in small perspex trays and incubated in 10 ml secondary antibody solution (antibody diluted in PBST) for one hour with gentle agitation. Routinely, the anti-rabbit secondary antibody supplied with the ECL kit (Amersham, cat.-no. NA934) was diluted 1:2000.
6. The membranes were briefly rinsed twice with PBST and washed in large volumes (approx. 50 ml) of PBST: 1 x 15 minutes, 3 x 5 minutes.

7. The membranes were washed in $dH_2O$ twice for 2 minutes.
8. Excess liquid was removed from the membranes. The membranes were placed on cling-film, covered with the detection mix (1:1 mixture of supplied solutions 1 and 2) and incubated for exactly one minutes
9. Excess detection reagent was removed and the membranes were wrapped in cling-film and placed in a film cassette.
10. The blot was exposed to Hyperfilm-ECL autoradiography film and subsequently developed. Exposure times varied from a few seconds to 10 minutes.

## 6.12.4 Immuno-affinity purification of anti-Cdc17-peptide antibody

A protein fusion of the catalytic domains of Cdc17 (amino acids 391-768) C-terminal to GST had been used to raise antibodies as part of this work. Although the antibodies were not employed for the results presented here, the GST-fusion was used to affinity-purify the anti-Cdc17-peptide antibody. The following section describes the production of GST-Cdc17-NΔ390 and after that the purification procedure is described.

### 6.12.4.1 Production of GST-Cdc17-NΔ390

The GST-fusion was created as follows. *cdc17-NΔ390* was subcloned from pTZ19R-cdc17-NΔ390 into the *BamHI* site of pGEX6P-1B to produce pGEX6P-1B-cdc17-NΔ390. Alongside pGEX6P-1B as a control this plasmid was transformed in *E.coli* BL21(DE3)pLysS. Cells were grown to an OD of 0.5 at 37°C, induced with 0.1 mM IPTG and harvested after 3 hours. Since it was only possible to obtain very little soluble fusion protein a protocol for the purification of inclusion bodies was used (Yasuhisa Adachi, personal communication). Harvested cells were frozen, thawed, resuspended in TBS,1 mM β-mercaptoethanol, 0.2 mM PMSF and sonicated at 10% for 20 seconds with a high intensity ultrasonic processor (Sonics & Materials). The cells were spun down and resuspended in TBS,1 mM β-mercaptoethanol, 0.2 mM

PMSF, 2M urea, 2% Triton-X100 and sonicated at 5% for 10 seconds. The cells were pelleted, washed in same buffer and resuspended in TBS, 1 mM β-mercaptoethanol. This crude protein extract was mixed with one volume of SDS sample buffer and run on a small preparative gel. The gel was stained with Coomassie Blue in 20% methanol, destained in 10% methanol and the protein band corresponding to GST-Cdc17-NΔ390 was excised. The gel was squashed, transferred to a microfuge tube, 1 ml of Protein Elution Buffer was added (15 mM $NH_4HCO_3$, 0.025% SDS, 1 mM DTT, 0.1 mM PMSF) and incubated overnight on a rotating wheel at room temperature. The gel squash was washed once with fresh protein elution buffer. The supernatants were pooled, lyophilised and the protein concentration was determined according to Bradford (Bio-Rad Protein Assay, cat.-no. 500-0002, Bradford, 1976).

### 6.12.4.2 Purification

The purification procedure is based on a protocol described in Harlow and Lane, 1988.

Approximately 20 µg of GST-fusion protein was loaded on a mini preparative gel (SDS-PAGE) and run as normally. The proteins were blotted onto a PVDF membrane as described previously. The blot was washed in PBS for 5 minutes, then incubated in Ponceau S (Sigma) for 5 minutes. After destaining for a maximum of a few seconds, the visible protein band was excised and the membrane strip was blocked in milk for one hour. After washing in PBS, the strip was cut into pieces and transferred into a microfuge tube containing 1 ml of the antiserum (final bleed). The tube was incubated overnight at 4°C on a rotating wheel. The membrane pieces were washed 5x in PBS followed by elution of the antibody from the membrane: 500 µl of 1M NaSCN were added to the tube and incubated on a wheel for 10 minutes at room temperature. The solution was neutralised by the addition of 5 ml PBS. Finally, sodium azide was added to a final concentration of 0.02%. The purified antibody was stored at 4°C and is usable for up to two months after purification.

# References


**Alfa, C., P.A. Fantes, J. Hyams, M. McLeod, and E. Warbrick**. 1993. Experiments with fission yeast : a laboratory course manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

**Almendral, J.M., D. Huebsch, P.A. Blundell, H. Macdonald-Bravo, and R. Bravo**. 1987. Cloning and sequence of the human nuclear protein cyclin: homology with DNA-binding proteins. Proc.Natl.Acad.Sci.U.S.A. **84**:1575-1579.

**Altschul, S.F. and E.V. Koonin**. 1998. Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. Trends.Biochem.Sci. **23**:444-447.

**Andaluz, E., G. Larriba, and R. Calderone**. 1996. A *Candida albicans* gene encoding a DNA ligase. Yeast. **12**:893-898.

**Andaluz, E., A. Ciudad, C.J. Rubio, R. Calderone, and G. Larriba**. 1999. Cell cycle regulation of a DNA ligase-encoding gene (*Ca*LIG4) from *Candida albicans*. Yeast. **15**:1199-1210.

**Aparicio, O.M., D.M. Weinstein, and S.P. Bell**. 1997. Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. Cell **91** :59-69.

**Araujo, S.J., F. Tirode, F. Coin, H. Pospiech, J.E. Syvaoja, M. Stucki, U. Hubscher, J.M. Egly, and R.D. Wood**. 2000. Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. Genes Dev. **14**:349-359.

**Aravind, L. and E.V. Koonin**. 1999. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. J.Mol.Biol. **287**:1023-1040.

**Bae, S.H. and Y.S. Seo**. 2000. Characterization of the enzymatic properties of the yeast dna2 Helicase/endonuclease suggests a new model for Okazaki fragment processing. J.Biol.Chem. **275**:38022-38031.

**Bae, S.H., K.H. Bae, J.A. Kim, and Y.S. Seo**. 2001. RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. Nature **412**:456-461.

**Bahler, J., J.Q. Wu, M.S. Longtine, N.G. Shah, A. McKenzie, A.B. Steever, A. Wach, P. Philippsen, and J.R. Pringle**. 1998. Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. Yeast. **14**:943-951.

**Baim, S.B. and F. Sherman**. 1988. mRNA structures influencing translation in the yeast *Saccharomyces cerevisiae*. Mol.Cell Biol. **8**:1591-1601.

**Barbet, N., W.J. Muriel, and A.M. Carr**. 1992. Versatile shuttle vectors and genomic libraries for use with *Schizosaccharomyces pombe*. Gene **114**:59-66.

**Barker, D.G. and L.H. Johnston**. 1983. *Saccharomyces cerevisiae cdc9*, a structural gene for yeast DNA ligase which complements *Schizosaccharomyces pombe cdc17*. Eur.J.Biochem. **134**:315-319.

**Barker, D.G., J.H. White, and L.H. Johnston**. 1985. The nucleotide sequence of the DNA ligase gene (*CDC9*) from *Saccharomyces cerevisiae*: a gene which is cell-cycle regulated and induced in response to DNA damage. Nucleic.Acids.Res. **13**:8323-8337.

**Barker, D.G., J.H. White, and L.H. Johnston**. 1987. Molecular characterisation of the DNA ligase gene, *CDC17*, from the fission yeast *Schizosaccharomyces pombe*. Eur.J.Biochem. **162**:659-667.

**Barnes, D.E., L.H. Johnston, K. Kodama, A.E. Tomkinson, D.D. Lasko, and T. Lindahl**. 1990. Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae*. Proc.Natl.Acad.Sci.U.S.A. **87**:6679-6683.

**Barnes, D.E., A.E. Tomkinson, A.R. Lehmann, A.D. Webster, and T. Lindahl**. 1992. Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. Cell **69**:495-503.

**Barnes, D.E., G. Stamp, I. Rosewell, A. Denzel, and T. Lindahl**. 1998. Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. Curr.Biol. **8**:1395-1398.

**Basi, G., E. Schmid, and K. Maundrell**. 1993. TATA box mutations in the *Schizosaccharomyces pombe nmt1* promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. Gene **123**:131-136.

**Bauer, G.A. and P.M. Burgers**. 1988a. The yeast analog of mammalian cyclin/proliferating-cell nuclear antigen interacts with mammalian DNA polymerase delta. Proc.Natl.Acad.Sci.U.S.A. **85**:7506-7510.

**Bauer, G.A. and P.M. Burgers**. 1988b. Protein-protein interactions of yeast DNA polymerase III with mammalian and yeast proliferating cell nuclear antigen (PCNA)/cyclin. Biochim.Biophys.Acta **951**:274-279.

**Bauer, G.A. and P.M. Burgers**. 1990. Molecular cloning, structure and expression of the yeast proliferating cell nuclear antigen gene. Nucleic.Acids.Res. **18**:261-265.

**Bell, S.P. and B. Stillman**. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature **357**:128-134.

**Bentley, D., J. Selfridge, J.K. Millar, K. Samuel, N. Hole, J.D. Ansell, and D.W. Melton**. 1996. DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability. Nat.Genet. **13**:489-491.

**Bielinsky, A.K. and S.A. Gerbi**. 1999. Chromosomal *ARS*1 has a single leading strand start site. Mol.Cell **3**:477-486.

**Bogenhagen, D.F., K.G. Pinz, and R.M. Perez-Jannotti**. 1902. Enzymology of mitochondrial base excision repair. Prog.Nucleic.Acid.Res.Mol.Biol. **68**:257-271.

**Bostock, C.J.** 1969. Mitochondrial DNA in the fission yeast *Schizosaccharomyces pombe*. Biochim.Biophys.Acta **195**:579-581.

**Bowman, K.K., K. Sidik, C.A. Smith, J.S. Taylor, P.W. Doetsch, and G.A. Freyer**. 1994. A new ATP-independent DNA endonuclease from *Schizosaccharomyces pombe* that recognizes cyclobutane pyrimidine dimers and 6-4 photoproducts. Nucleic.Acids.Res. **22**:3026-3032.

**Bravo, R. and J.E. Celis**. 1980. A search for differential polypeptide synthesis throughout the cell cycle of *HeLa* cells. J.Cell Biol. **84**:795-802.

**Budd, M.E. and J.L. Campbell**. 1993. DNA polymerases delta and epsilon are required for chromosomal replication in *Saccharomyces cerevisiae*. Mol.Cell Biol. **13**:496-505.

**Budd, M.E. and J.L. Campbell**. 1995. A yeast gene required for DNA replication encodes a protein with homology to DNA helicases. Proc.Natl.Acad.Sci.U.S.A. **92**:7642-7646.

**Budd, M.E. and J.L. Campbell**. 1997. A yeast replicative helicase, Dna2 helicase, interacts with yeast FEN-1 nuclease in carrying out its essential function. Mol.Cell Biol. **17**:2136-2142.

**Budd, M.E., W. Choe, and J.L. Campbell**. 2000. The nuclease activity of the yeast DNA2 protein, which is related to the RecB-like nucleases, is essential *in vivo*. J.Biol.Chem. **275**:16518-16529.

**Burgers, P.M. and K.J. Gerik**. 1998. Structure and processivity of two forms of *Saccharomyces cerevisiae* DNA polymerase delta. J.Biol.Chem. **273**:19756-19762.

**Caldecott, K.W., C.K. McKeown, J.D. Tucker, S. Ljungquist, and L.H. Thompson**. 1994. An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. Mol.Cell Biol. **14**:68-76.

**Caldecott, K.W., J.D. Tucker, L.H. Stanker, and L.H. Thompson**. 1995. Characterization of the XRCC1-DNA ligase III complex *in vitro* and its absence from mutant hamster cells. Nucleic.Acids.Res. **23**:4836-4843.

**Caldecott, K.W., S. Aoufouchi, P. Johnson, and S. Shall**. 1996. XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' *in vitro*. Nucleic.Acids.Res. **24**:4387-4394.

**Cappelli, E., R. Taylor, M. Cevasco, A. Abbondandolo, K. Caldecott, and G. Frosina**. 1997. Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. J.Biol.Chem. **272** :23970-23975.

**Cardoso, M.C., C. Joseph, H.P. Rahn, R. Reusch, B. Nadal-Ginard, and H. Leonhardt**. 1997. Mapping and use of a sequence that targets DNA ligase I to sites of DNA replication *in vivo*. J.Cell Biol. **139**:579-587.

**Chen, J., A.E. Tomkinson, W. Ramos, Z.B. Mackey, S. Danehower, C.A. Walter, R.A. Schultz, J.M. Besterman, and I. Husain**. 1995. Mammalian DNA ligase III: molecular cloning, chromosomal localization, and expression in spermatocytes undergoing meiotic recombination. Mol.Cell Biol. **15**:5412-5422.

**Cheng, C. and S. Shuman**. 1997. Characterization of an ATP-dependent DNA ligase encoded by *Haemophilus influenzae*. Nucleic.Acids.Res. **25**:1369-1374.

**Christophe, D., C. Christophe-Hobertus, and B. Pichon**. 2000. Nuclear targeting of proteins: how many different signals? Cell Signal. **12**:337-341.

**Claros, M.G. and P. Vincens**. 1996. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur.J.Biochem. **241**:779-786.

**Claros, M.G., S. Brunak, and G. von Heijne**. 1997. Prediction of N-terminal protein sorting signals. Curr.Opin.Struct.Biol. 7:394-398.

**Clayton, D.A.** 1982. Replication of animal mitochondrial DNA. Cell **28**:693-705.

**Clayton, D.A.** 1996. Mitochondrial DNA replication. p. 1015-1027. In M.L. DePamphilis. (ed.), DNA replication in eukaryotic cells. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

**Critchlow, S.E., R.P. Bowater, and S.P. Jackson**. 1997. Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. Curr.Biol. **7**:588-598.

**Cullmann, G., K. Fien, R. Kobayashi, and B. Stillman**. 1995. Characterization of the five replication factor C genes of *Saccharomyces cerevisiae*. Mol.Cell Biol. **15**:4661-4671.

**D'Urso, G. and P. Nurse**. 1997. *Schizosaccharomyces pombe cdc20+* encodes DNA polymerase epsilon and is required for chromosomal replication but not for the S phase checkpoint. Proc.Natl.Acad.Sci.U.S.A. **94**:12491-12496.

**Danpure, C.J.** 1995. How can the products of a single gene be localized to more than one intracellular compartment? Trends Cell Biol. **5**:230-238.

**Davey, M.J. and M. O'Donnell**. 2000. Mechanisms of DNA replication. Curr.Opin.Chem.Biol. **4**:581-586.

**Dimitriadis, E.K., R. Prasad, M.K. Vaske, L. Chen, A.E. Tomkinson, M.S. Lewis, and S.H. Wilson**. 1998. Thermodynamics of human DNA ligase I trimerization and association with DNA polymerase beta. J.Biol.Chem. **273**:20540-20550.

**Ding, D.Q., Y. Tomita, A. Yamamoto, Y. Chikashige, T. Haraguchi, and Y. Hiraoka**. 2000. Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a GFP-fusion genomic DNA library. Genes Cells **5**:169-190.

**Doe, C.L., F. Osman, J. Dixon, and M.C. Whitby**. 2000. The Holliday junction resolvase *SpCCE1* prevents mitochondrial DNA aggregation in *Schizosaccharomyces pombe*. Mol.Gen.Genet. **263**:889-897.

**Doherty, A.J. and S.W. Suh**. 2000. Structural and mechanistic conservation in DNA ligases. Nucleic.Acids.Res. **28**:4051-4058.

**Donahue, S.L., B.E. Corner, L. Bordone, and C. Campbell**. 2001. Mitochondrial DNA ligase function in *Saccharomyces cerevisiae*. Nucleic.Acids.Res. **29**:1582-1589.

**Dua, R., D.L. Levy, and J.L. Campbell**. 1999. Analysis of the essential functions of the C-terminal protein/protein interaction domain of *Saccharomyces cerevisiae* pol epsilon and its unexpected ability to support growth in the absence of the DNA polymerase domain. J.Biol.Chem. **274**:22283-22288.

**Ducoux, M., S. Urbach, G. Baldacci, U. Hubscher, S. Koundrioukoff, J. Christensen, and P. Hughes**. 2001. Mediation of PCNA-dependent DNA replication through a conserved P21Cip1-like PCNA-binding motif present in the third subunit of human DNA polymerase delta. J.Biol.Chem.

**Dutta, A. and S.P. Bell**. 1997. Initiation of DNA replication in eukaryotic cells. Annu.Rev.Cell Dev.Biol. **13**:293-332.

**Edenberg, H.J. and J.A. Huberman**. 1975. Eukaryotic chromosome replication. Annu.Rev.Genet. 9:245-284.

**Eki, T., K. Okumura, A. Shiratori, M. Abe, M. Nogami, H. Taguchi, T. Shibata, Y. Murakami, and F. Hanaoka**. 1996. Assignment of the closest human homologue (DNA2L:KIAA0083) of the yeast Dna2 helicase gene to chromosome band 10q21.3-q22.1. Genomics **37**:408-410.

**Fiorentino, D.F. and G.R. Crabtree**. 1997. Characterization of *Saccharomyces cerevisiae* dna2 mutants suggests a role for the helicase late in S phase. Mol.Biol.Cell **8**:2519-2537.

**Forsburg, S.L.** 1993. Comparison of *Schizosaccharomyces pombe* expression systems. Nucleic.Acids.Res. **21**:2955-2956.

**Fotedar, R., R. Mossi, P. Fitzgerald, T. Rousselle, G. Maga, H. Brickner, H. Messier, S. Kasibhatla, U. Hubscher, and A. Fotedar**. 1996. A conserved domain of the large subunit of replication factor C binds PCNA and acts like a dominant negative inhibitor of DNA replication in mammalian cells. EMBO J. **15**:4423-4433.

**Frank, G., J. Qiu, L. Zheng, and B. Shen**. 2001. Stimulation of eukaryotic flap endonuclease-1 activities by proliferating cell nuclear antigen (PCNA) is independent of its *in vitro* interaction via a consensus PCNA binding region. J.Biol.Chem. **276**:36295-36302.

**Frank, K.M., J.M. Sekiguchi, K.J. Seidl, W. Swat, G.A. Rathbun, H.L. Cheng, L. Davidson, L. Kangaloo, and F.W. Alt**. 1998. Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. Nature **396**:173-177.

**Gao, Y., Y. Sun, K.M. Frank, P. Dikkes, Y. Fujiwara, K.J. Seidl, J.M. Sekiguchi, G.A. Rathbun, W. Swat, J. Wang, R.T. Bronson, B.A. Malynn, M. Bryans, C. Zhu, J. Chaudhuri, L. Davidson, R. Ferrini, T. Stamato, S.H. Orkin, M.E. Greenberg, and F.W. Alt**. 1998. A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. Cell **95**:891-902.

**Gary, R., D.L. Ludwig, H.L. Cornelius, M.A. MacInnes, and M.S. Park**. 1997. The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence

elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21. J.Biol.Chem. **272**:24522-24529.

**Gavin, K.A., M. Hidaka, and B. Stillman**. 1995. Conserved initiator proteins in eukaryotes. Science **270**:1667-1671.

**Gerik, K.J., X. Li, A. Pautz, and P.M. Burgers**. 1998. Characterization of the two small subunits of *Saccharomyces cerevisiae* DNA polymerase delta. J.Biol.Chem. **273**:19747-19755.

**Gomes, X.V., S.L. Gary, and P.M. Burgers**. 2000. Overproduction in *Escherichia coli* and characterization of yeast replication factor C lacking the ligase homology domain. J.Biol.Chem. **275**:14541-14549.

**Gomes, X.V. and P.M. Burgers**. 2000. Two modes of FEN1 binding to PCNA regulated by DNA. EMBO J. **19**:3811-3821.

**Gomez, M. and F. Antequera**. 1999. Organization of DNA replication origins in the fission yeast genome. EMBO J. **18**:5683-5690.

**Grawunder, U., M. Wilm, X. Wu, P. Kulesza, T.E. Wilson, M. Mann, and M.R. Lieber**. 1997. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature **388**:492-495.

**Grawunder, U., D. Zimmer, and M.R. Leiber**. 1998. DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains. Curr.Biol. **8**:873-876.

**Gulbis, J.M., Z. Kelman, J. Hurwitz, M. O'Donnell, and J. Kuriyan**. 1996. Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. Cell **87**:297-306.

**Hakansson, K., A.J. Doherty, S. Shuman, and D.B. Wigley**. 1997. X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. Cell **89**:545-553.

**Han, Z. and C. Stachow**. 1994. Analysis of *Schizosaccharomyces pombe* mitochondrial DNA replication by two dimensional gel electrophoresis. Chromosoma **103**:162-170.

**Harlow, E. and D. Lane**. 1988. Antibodies : a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

**Hartwell, L.H., J. Culotti, J.R. Pringle, and B.J. Reid**. 1974. Genetic control of the cell division cycle in yeast. Science **183**:46-51.

**Hassell, J.A. and B.T. Brinton**. 1996. SV40 and polyomavirus DNA replication. p. 639-677. In M.L. DePamphilis. (ed.), DNA replication in eukaryotic cells. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

**Heim, R., A.B. Cubitt, and R.Y. Tsien**. 1995. Improved green fluorescence. Nature **373**:663-664.

**Herrmann, G., T. Lindahl, and P. Schar**. 1998. *Saccharomyces cerevisiae* LIF1: a function involved in DNA double-strand break repair related to mammalian XRCC4. EMBO J. **17**:4188-4198.

**Holt, I.J., H.E. Lorimer, and H.T. Jacobs**. 2000. Coupled leading- and lagging strand synthesis of mammalian mitochondrial DNA. Cell **100**:515-524.

**Hughes, P., I. Tratner, M. Ducoux, K. Piard, and G. Baldacci**. 1999. Isolation and identification of the third subunit of mammalian DNA polymerase delta by PCNA-affinity chromatography of mouse FM3A cell extracts. Nucleic.Acids.Res. **27**:2108-2114.

**Husain, I., A.E. Tomkinson, W.A. Burkhart, M.B. Moyer, W. Ramos, Z.B. Mackey, J.M. Besterman, and J. Chen**. 1995. Purification and characterization of DNA ligase III from bovine testes. Homology with DNA ligase II and vaccinia DNA ligase. J.Biol.Chem. **270**:9683-9690.

**Ikejima, M., S. Noguchi, R. Yamashita, T. Ogura, T. Sugimura, D.M. Gill, and M. Miwa**. 1990. The zinc fingers of human poly(ADP-ribose) polymerase are differentially required for the recognition of DNA breaks and nicks and the consequent enzyme activation. Other structures recognize intact DNA. J.Biol.Chem. **265**:21907-21913.

**Ireland, M.J., S.S. Reinke, and D.M. Livingston**. 2000. The impact of lagging strand replication mutations on the stability of CAG repeat tracts in yeast. Genetics **155**:1657-1665.

**Ishimi, Y.** 1997. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. J.Biol.Chem. **272**:24508-24513.

**Jaskulski, D., J.K. deRiel, W.E. Mercer, B. Calabretta, and R. Baserga**. 1988. Inhibition of cellular proliferation by antisense oligodeoxynucleotides to PCNA cyclin. Science **240**:1544-1546.

**Jeruzalmi, D., M. O'Donnell, and J. Kuriyan**. 2001. Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III. Cell **106**:429-441.

**Johnson, A.P. and M.P. Fairman**. 1997. The identification and purification of a novel mammalian DNA ligase. Mutat.Res. **383**:205-212.

**Johnston, L.H. and K.A. Nasmyth**. 1978. *Saccharomyces cerevisiae* cell cycle mutant *cdc9* is defective in DNA ligase. Nature **274**:891-893.

**Johnston, L.H., D.G. Barker, and P. Nurse**. 1986. Cloning and characterization of the *Schizosaccharomyces pombe* DNA ligase gene *CDC17*. Gene **41**:321-325.

**Jonsson, Z.O., R. Hindges, and U. Hubscher**. 1998. Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. EMBO J. **17** :2412-2425.

**Kang, H.Y., E. Choi, S.H. Bae, K.H. Lee, B.S. Gim, H.D. Kim, C. Park, S.A. MacNeill, and Y.S. Seo**. 2000. Genetic analyses of *Schizosaccharomyces pombe dna2(+)* reveal that *dna2* plays an essential role in Okazaki fragment metabolism. Genetics **155**:1055-1067.

**Karthikeyan, R., E.J. Vonarx, A.F. Straffon, M. Simon, G. Faye, and B.A. Kunz**. 2000. Evidence from mutational specificity studies that yeast DNA polymerases delta and epsilon replicate different DNA strands at an intracellular replication fork. J.Mol.Biol. **299**:405-419.

**Kelman, Z. and M. O'Donnell**. 1995. Structural and functional similarities of prokaryotic and eukaryotic DNA polymerase sliding clamps. Nucleic.Acids.Res. **23**:3613-3620.

**Kelman, Z.** 1997. PCNA: structure, functions and interactions. Oncogene **14**:629-640.

**Kesti, T., K. Flick, S. Keranen, J.E. Syvaoja, and C. Wittenberg**. 1999. DNA polymerase epsilon catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. Mol.Cell **3**:679-685.

**Kleczkowska, H.E., G. Marra, T. Lettieri, and J. Jiricny** . 2001. hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. Genes Dev. **15**:724-736.

**Kletzin, A.** 1992. Molecular characterisation of a DNA ligase gene of the extremely thermophilic archaeon *Desulfurolobus ambivalens* shows close phylogenetic relationship to eukaryotic ligases. Nucleic.Acids.Res. **20**:5389-5396.

**Kodama, K., D.E. Barnes, and T. Lindahl**. 1991. *In vitro* mutagenesis and functional expression in *Escherichia coli* of a cDNA encoding the catalytic domain of human DNA ligase I. Nucleic.Acids.Res. **19**:6093-6099.

**Kornberg, A. and T.A. Baker**. 1992. DNA replication. Freeman, New York.

**Koundrioukoff, S., Z.O. Jonsson, S. Hasan, R.N. de Jong, d. van, V, M.O. Hottiger, and U. Hubscher**. 2000. A direct interaction between proliferating cell nuclear antigen (PCNA) and

Cdk2 targets PCNA-interacting proteins for phosphorylation. J.Biol.Chem. **275**:22882-22887.

**Kozak, M.** 1996. Interpreting cDNA sequences: some insights from studies on translation. Mamm.Genome 7:563-574.

**Krishna, T.S., X.P. Kong, S. Gary, P.M. Burgers, and J. Kuriyan**. 1994. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell **79**:1233-1243.

**Kubota, Y., R.A. Nash, A. Klungland, P. Schar, D.E. Barnes, and T. Lindahl**. 1996. Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. EMBO J. **15**:6662-6670.

**Kunkel, T.A., J.D. Roberts, and R.A. Zakour**. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367-382.

**Labib, K., J.A. Tercero, and J.F. Diffley**. 2000. Uninterrupted MCM2-7 function required for DNA replication fork progression. Science **288**:1643-1647.

**Labib, K. and J.F. Diffley**. 2001. Is the MCM2-7 complex the eukaryotic DNA replication fork helicase? Curr.Opin.Genet.Dev. **11**:64-70.

**Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

**Lakshmipathy, U. and C. Campbell**. 1999. The human DNA ligase III gene encodes nuclear and mitochondrial proteins. Mol.Cell Biol. **19**:3869-3876.

**Lakshmipathy, U. and C. Campbell**. 2000. Mitochondrial DNA ligase III function is independent of Xrcc1. Nucleic.Acids.Res. **28**:3880-3886.

**Lakshmipathy, U. and C. Campbell**. 2001. Antisense-mediated decrease in DNA ligase III expression results in reduced mitochondrial DNA integrity. Nucleic.Acids.Res. **29**:668-676.

**Lasko, D.D., A.E. Tomkinson, and T. Lindahl**. 1990a. Mammalian DNA ligases. Biosynthesis and intracellular localization of DNA ligase I. J.Biol.Chem. **265**:12618-12622.

**Lasko, D.D., A.E. Tomkinson, and T. Lindahl**. 1990b. Eukaryotic DNA ligases. Mutat.Res. **236**:277-287.

**Lee, J., P.D. Chastain, T. Kusakabe, J.D. Griffith, and C.C. Richardson**. 1998. Coordinated leading and lagging strand DNA synthesis on a minicircular template. Mol Cell **1**:1001-1010.

**Lee, K.H., D.W. Kim, S.H. Bae, J.A. Kim, G.H. Ryu, Y.N. Kwon, K.A. Kim, H.S. Koo, and Y.S. Seo**. 2000a. The endonuclease activity of the yeast Dna2 enzyme is essential *in vivo*. Nucleic.Acids.Res. **28**:2873-2881.

**Lee, J.Y., C. Chang, H.K. Song, J. Moon, J.K. Yang, H.K. Kim, S.T. Kwon, and S.W. Suh**. 2000b. Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications. EMBO J. **19**:1119-1129.

**Lee, S.H., A.D. Kwong, Z.Q. Pan, and J. Hurwitz**. 1991a. Studies on the activator 1 protein complex, an accessory factor for proliferating cell nuclear antigen-dependent DNA polymerase delta. J.Biol.Chem. **266**:594-602.

**Lee, S.H., Z.Q. Pan, A.D. Kwong, P.M. Burgers, and J. Hurwitz**. 1991b. Synthesis of DNA by DNA polymerase epsilon *in vitro*. J.Biol.Chem. **266**:22707-22717.

**Lehman, I.R.** 1974. DNA ligase: structure, mechanism, and function. Science **186**:790-797.

**Lei, M. and B.K. Tye**. 2001. Initiating DNA synthesis: from recruiting to activating the MCM complex. J.Cell Sci. **114**:1447-1454.

**Lengronne, A., P. Pasero, A. Bensimon, and E. Schwob**. 2001. Monitoring S phase progression globally and locally using BrdU incorporation in TK(+) yeast strains. Nucleic.Acids.Res. **29**:1433-1442.

**Leonhardt, H., A.W. Page, H.U. Weier, and T.H. Bestor**. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell **71**:865-873.

**Leonhardt, H., A. Sporbert, and M.C. Cardoso**. 2000. Targeting regulatory factors to intranuclear replication sites. Crit.Rev.Eukaryot.Gene Expr. **10**:127-133.

**Lepetit, D., P. Thiebaud, S. Aoufouchi, C. Prigent, R. Guesne, and N. Theze**. 1996. The cloning and characterization of a cDNA encoding *Xenopus laevis* DNA ligase I. Gene **172**:273-277.

**Levin, C.J. and S.B. Zimmerman**. 1976. A DNA ligase from mitochondria of rat liver. Biochem.Biophys.Res.Commun. **69**:514-520.

**Levin, D.S., W. Bai, N. Yao, M. O'Donnell, and A.E. Tomkinson**. 1997. An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. Proc.Natl.Acad.Sci.U.S.A. **94**:12863-12868.

**Levin, D.S., A.E. McKenna, T.A. Motycka, Y. Matsumoto, and A.E. Tomkinson**. 2000. Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. Curr.Biol. **10**:919-922.

**Li, C., J. Goodchild, and E.F. Baril**. 1994. DNA ligase I is associated with the 21 S complex of enzymes for DNA synthesis in *HeLa* cells. Nucleic.Acids.Res. **22**:632-638.

**Li, X., J. Li, J. Harrington, M.R. Lieber, and P.M. Burgers**. 1995. Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. J.Biol.Chem. **270**:22109-22112.

**Liang, C., M. Weinreich, and B. Stillman**. 1995. ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. Cell **81**:667-676.

**Liang, C. and B. Stillman**. 1997. Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in cdc6 mutants. Genes Dev. **11**:3375-3386.

**Lieber, M.R.** 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. Bioessays **19**:233-240.

**Lindahl, T. and G.M. Edelman**. 1968. Polynucleotide ligase from myeloid and lymphoid tissues. Proc.Natl.Acad.Sci.U.S.A. **61**:680-687.

**Lithgow, T.** 2000. Targeting of proteins to mitochondria. FEBS Lett. **476**:22-26.

**Liu, Q., W. Choe, and J.L. Campbell**. 2000. Identification of the *Xenopus laevis* homolog of *Saccharomyces cerevisiae* DNA2 and its role in DNA replication. J.Biol.Chem. **275**:1615-1624.

**Ljungquist, S., K. Kenne, L. Olsson, and M. Sandstrom**. 1994. Altered DNA ligase III activity in the CHO EM9 mutant. Mutat.Res. **314**:177-186.

**Loor, G., S.J. Zhang, P. Zhang, N.L. Toomey, and M.Y. Lee**. 1997. Identification of DNA replication and cell cycle proteins that interact with PCNA. Nucleic.Acids.Res. **25**:5041-5046.

**Luo, J. and F. Barany**. 1996. Identification of essential residues in *Thermus thermophilus* DNA ligase. Nucleic.Acids.Res. **24**:3079-3085.

**Mackenney, V.J., D.E. Barnes, and T. Lindahl**. 1997. Specific function of DNA ligase I in simian virus 40 DNA replication by human cell-free extracts is mediated by the amino-terminal non-catalytic domain. J.Biol.Chem. **272**:11550-11556.

**Mackey, Z.B., W. Ramos, D.S. Levin, C.A. Walter, J.R. McCarrey, and A.E. Tomkinson**. 1997. An alternative splicing event which occurs in mouse pachytene spermatocytes generates a form of DNA ligase III with distinct biochemical properties that may function in meiotic recombination. Mol.Cell Biol. **17**:989-998.

**Mackey, Z.B., C. Niedergang, J.M. Murcia, J. Leppard, K. Au, J. Chen, G. de Murcia, and A.E. Tomkinson**. 1999. DNA ligase III is recruited to DNA strand breaks by a zinc finger motif homologous to that of poly(ADP-ribose) polymerase. Identification of two functionally distinct DNA binding regions within DNA ligase III. J.Biol.Chem. **274**:21679-21687.

**MacNeill, S.A. and P.A. Fantes**. 1993. Methods for analysis of the fission yeast cell cycle. p. 93-125. In P. Fantes and R. Brooks. (ed.), The Cell cycle : a practical approach. Oxford University press, Oxford, UK.

**MacNeill, S.A., S. Moreno, N. Reynolds, P. Nurse, and P.A. Fantes**. 1996. The fission yeast Cdc1 protein, a homologue of the small subunit of DNA polymerase delta, binds to Pol3 and Cdc27. EMBO J. **15**:4613-4628.

**MacNeill, S.A. and P.A. Fantes**. 1997. Genetic and physiological analysis of DNA replication in fission yeast. Methods Enzymol. **283**:440-459.

**MacNeill, S.A. and P. Nurse**. 1997. Cell cycle control in fission yeast. p. Anonymous The Molecular and Cellular Biology of Yeast. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

**MacNeill, S.A. and P.M.J. Burgers**. 2000. Chromosomal DNA replication in yeast: enzymes and mechanisms. p. 19-57. In P. Fantes and J. Beggs. (ed.), The Yeast Nucleus. Oxford University Press, Oxford, UK.

**MacNeill, S.A.** 2001a. Understanding the enzymology of archaeal DNA replication: progress in form and function. Mol.Microbiol. **40**:520-529.

**MacNeill, S.A.** 2001b. DNA replication: partners in the Okazaki two-step. Curr.Biol. **11**:R842-R844

**MacNeill, S.A., G. Baldacci, P.M. Burgers, and U. Hubscher**. 2001. A unified nomenclature for the subunits of eukaryotic DNA polymerase delta. Trends.Biochem.Sci. **26**:16-17.

**Maiorano, D., J. Moreau, and M. Mechali**. 2000. XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis. Nature **404**:622-625.

**Manolis, K.G., E.R. Nimmo, E. Hartsuiker, A.M. Carr, P.A. Jeggo, and R.C. Allshire**. 2001. Novel functional requirements for non-homologous DNA end joining in *Schizosaccharomyces pombe*. EMBO J. **20**:210-221.

**Maundrell, K.** 1990. nmt1 of fission yeast. A highly transcribed gene completely repressed by thiamine. J.Biol.Chem. **265**:10857-10864.

**Maundrell, K.** 1993. Thiamine-repressible expression vectors pREP and pRIP for fission yeast. Gene **123**:127-130.

**Melendy, T. and B. Stillman**. 1991. Purification of DNA polymerase delta as an essential simian virus 40 DNA replication factor. J.Biol.Chem. **266**:1942-1949.

**Mitchison, J. M**. The growth of single cells in *Schizosaccharomyces pombe*. Exp.Cell Res. **13**, 244-262. 1957.

**Miyachi, K., M.J. Fritzler, and E.M. Tan**. 1978. Autoantibody to a nuclear antigen in proliferating cells. J.Immunol. **121**:2228-2234.

**Modesti, M., J.E. Hesse, and M. Gellert**. 1999. DNA binding of Xrcc4 protein is associated with V(D)J recombination but not with stimulation of DNA ligase IV activity. EMBO J. **18**:2008-2018.

**Montecucco, A., E. Savini, F. Weighardt, R. Rossi, G. Ciarrocchi, A. Villa, and G. Biamonti**. 1995. The N-terminal domain of human DNA ligase I contains the nuclear localization signal and directs the enzyme to sites of DNA replication. EMBO J. **14**:5379-5386.

**Montecucco, A., R. Rossi, D.S. Levin, R. Gary, M.S. Park, T.A. Motycka, G. Ciarrocchi, A. Villa, G. Biamonti, and A.E. Tomkinson**. 1998. DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories. EMBO J. **17** :3786-3795.

**Moon, K.Y., D. Kong, J.K. Lee, S. Raychaudhuri, and J. Hurwitz**. 1999. Identification and reconstitution of the origin recognition complex from *Schizosaccharomyces pombe*. Proc.Natl.Acad.Sci.U.S.A. **96**:12367-12372.

**Moreno, S., A. Klar, and P. Nurse**. 1991. Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. Methods Enzymol. **194**:795-823.

**Morrison, A. and A. Sugino**. 1994. The 3'-->5' exonucleases of both DNA polymerases delta and epsilon participate in correcting errors of DNA replication in *Saccharomyces cerevisiae*. Mol.Gen.Genet. **242**:289-296.

**Mossi, R., E. Ferrari, and U. Hubscher**. 1998. DNA ligase I selectively affects DNA synthesis by DNA polymerases delta and epsilon suggesting differential functions in DNA replication and repair. J.Biol.Chem. **273**:14322-14330.

**Murante, R.S., L.A. Henricksen, and R.A. Bambara**. 1998. Junction ribonuclease: an activity in Okazaki fragment processing. Proc.Natl.Acad.Sci.U.S.A. **95**:2244-2249.

**Nakai, K. and M. Kanehisa**. 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics **14**:897-911.

**Nakatani, M., S. Ezaki, H. Atomi, and T. Imanaka**. 2000. A DNA ligase from a hyperthermophilic archaeon with unique cofactor specificity. J.Bacteriol. **182**:6424-6433.

**Nash, R.A., K.W. Caldecott, D.E. Barnes, and T. Lindahl**. 1997. XRCC1 protein interacts with one of two distinct forms of DNA ligase III. Biochemistry **36**:5207-5211.

**Nasmyth, K. and P. Nurse**. 1981. Cell division cycle mutants altered in DNA replication and mitosis in the fission yeast *Schizosaccharomyces pombe*. Mol.Gen.Genet. **182**:119-124.

**Nasmyth, K.A.** 1977. Temperature-sensitive lethal mutants in the structural gene for DNA ligase in the yeast *Schizosaccharomyces pombe*. Cell **12**:1109-1120.

**Nasmyth, K.A.** 1979. Genetic and enzymatic characterization of conditional lethal mutants of the yeast *Schizosaccharomyces pombe* with a temperature-sensitive DNA ligase. J.Mol.Biol. **130**:273-284.

**Neupert, W.** 1997. Protein import into mitochondria. Annu.Rev.Biochem. **66**:863-917.

**Newman, J. and P. Hanawalt**. 1968. Intermediates in T4 DNA replication in a T4 ligase deficient strain. Cold Spring Harb.Symp.Quant.Biol. **33**:145-150.

**Nigg, E.A.** 1997. Nucleocytoplasmic transport: signals, mechanisms and regulation. Nature **386**:779-787.

**Niimi, A., N. Suka, M. Harata, A. Kikuchi, and S. Mizuno** . 2001. Co-localization of chicken DNA topoisomerase IIalpha, but not beta, with sites of DNA replication and possible involvement of a C-terminal region of alpha through its binding to PCNA. Chromosoma **110**:102-114.

**Nilsen, H., M. Otterlei, T. Haug, K. Solum, T.A. Nagelhus, F. Skorpen, and H.E. Krokan**. 1997. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the *UNG* gene. Nucleic.Acids.Res. **25**:750-755.

**Nishitani, H., Z. Lygerou, T. Nishimoto, and P. Nurse**. 2000. The Cdt1 protein is required to license DNA for replication in fission yeast. Nature **404**:625-628.

**Nurse, P.** 1975. Genetic control of cell size at cell division in yeast. Nature **256**:547-551.

**Nurse, P., P. Thuriaux, and K. Nasmyth**. 1976. Genetic control of the cell division cycle in the fission yeast *Schizosaccharomyces pombe*. Mol.Gen.Genet. **146**:167-178.

**Nurse, P. and P. Thuriaux**. 1977. Controls over the timing of DNA replication during the cell cycle of fission yeast. Exp.Cell Res. **107**:365-375.

**Nurse, P. and Y. Bissett**. 1981. Gene required in G1 for commitment to cell cycle and in G2 for control of mitosis in fission yeast. Nature **292**:558-560.

**O'Donnell, M., D. Jeruzalmi, and J. Kuriyan**. 2001. Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. Curr.Biol. **11**:R935-R946

**Odell, M., V. Sriskanda, V, S. Shuman, and D.B. Nikolov**. 2000. Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining. Mol.Cell **6**:1183-1193.

**Okazaki, R., T. Okazaki, K. Sakabe, K. Sugimoto, and A. Sugino**. 1968. Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. Proc.Natl.Acad.Sci.U.S.A. **59**:598-605.

**Otterlei, M., T. Haug, T.A. Nagelhus, G. Slupphaug, T. Lindmo, and H.E. Krokan**. 1998. Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively. Nucleic.Acids.Res. **26**:4611-4617.

**Oyama, T., Y. Ishino, I.K. Cann, S. Ishino, and K. Morikawa**. 2001. Atomic structure of the clamp loader small subunit from *Pyrococcus furiosus*. Mol.Cell **8**:455-463.

**Paquin, B., M.J. Laforest, L. Forget, I. Roewer, Z. Wang, J. Longcore, and B.F. Lang**. 1997. The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. Curr. Genet. **31**:380-395.

**Park, U.E., B.M. Olivera, K.T. Hughes, J.R. Roth, and D.R. Hillyard**. 1989. DNA ligase and the pyridine nucleotide cycle in *Salmonella typhimurium*. J.Bacteriol. **171**:2173-2180.

**Pasion, S.G. and S.L. Forsburg**. 1999. Nuclear localization of *Schizosaccharomyces pombe* Mcm2/Cdc19p requires MCM complex assembly. Mol.Biol.Cell **10**:4043-4057.

**Pelloquin, L., P. Belenguer, Y. Menon, N. Gas, and B. Ducommun**. 1999. Fission yeast Msp1 is a mitochondrial dynamin-related protein. J.Cell Sci. **112 ( Pt 22)**:4151-4161.

**Perez-Jannotti, R.M., S.M. Klein, and D.F. Bogenhagen**. 2001. Two forms of mitochondrial DNA ligase III are produced in *Xenopus laevis* oocytes. J.Biol.Chem. **276**:48978-48987.

**Petrini, J.H., Y. Xiao, and D.T. Weaver**. 1995. DNA ligase I mediates essential functions in mammalian cells. Mol.Cell Biol. **15**:4303-4308.

**Petrini, J.H., M.E. Walsh, C. DiMare, X.N. Chen, J.R. Korenberg, and D.T. Weaver**. 1995. Isolation and characterization of the human *MRE11* homologue. Genomics **29**:80-86.

**Pfanner, N., E.A. Craig, and A. Honlinger**. 1997. Mitochondrial preprotein translocase. Annu.Rev.Cell Dev.Biol. **13**:25-51.

**Pfanner, N.** 2000. Protein sorting: recognizing mitochondrial presequences. Curr.Biol. **10**:R412-R415

**Pinz, K.G. and D.F. Bogenhagen**. 1998. Efficient repair of abasic sites in DNA by mitochondrial enzymes. Mol.Cell Biol. **18**:1257-1265.

**Prasad, R., R.K. Singhal, D.K. Srivastava, J.T. Molina, A.E. Tomkinson, and S.H. Wilson**. 1996. Specific interaction of DNA polymerase beta and DNA ligase I in a multiprotein base excision repair complex from bovine testis. J.Biol.Chem. **271**:16000-16007.

**Prelich, G., M. Kostura, D.R. Marshak, M.B. Mathews, and B. Stillman**. 1987a. The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication *in vitro*. Nature **326**:471-475.

**Prelich, G., C.K. Tan, M. Kostura, M.B. Mathews, A.G. So, K.M. Downey, and B. Stillman**. 1987b. Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein. Nature **326**:517-520.

**Prentice, H.L.** 1992. High efficiency transformation of *Schizosaccharomyces pombe* by electroporation. Nucleic.Acids.Res. **20**:621

**Prigent, C., D.D. Lasko, K. Kodama, J.R. Woodgett, and T. Lindahl**. 1992. Activation of mammalian DNA ligase I through phosphorylation by casein kinase II. EMBO J. **11**:2925-2933.

**Prigent, C., M.S. Satoh, G. Daly, D.E. Barnes, and T. Lindahl**. 1994. Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I. Mol.Cell Biol. **14**:310-317.

**Qiu, J., Y. Qian, P. Frank, U. Wintersberger, and B. Shen**. 1999. *Saccharomyces cerevisiae* RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. Mol.Cell Biol. **19**:8361-8371.

**Ramos, W., G. Liu, C.N. Giroux, and A.E. Tomkinson**. 1998. Biochemical and genetic characterization of the DNA ligase encoded by *Saccharomyces cerevisiae* open reading frame YOR005c, a homolog of mammalian DNA ligase IV. Nucleic.Acids.Res. **26**:5676-5683.

**Ranalli, T.A., M.S. DeMott, and R.A. Bambara**. 2002. Mechanism underlying replication protein a stimulation of DNA ligase I. J. Biol. Chem. **277**:1719-1727.

**Reynolds, N., P.A. Fantes, and S.A. MacNeill**. 1999. A key role for replication factor C in DNA replication checkpoint function in fission yeast. Nucleic.Acids.Res. **27**:462-469.

**Reynolds, N., E. Warbrick, P.A. Fantes, and S.A. MacNeill**. 2000. Essential interaction between the fission yeast DNA polymerase delta subunit Cdc27 and Pcn1 (PCNA) mediated through a C-terminal p21(Cip1)-like PCNA binding motif. EMBO J. **19**:1108-1118.

**Robinow, C.F. and J.S. Hyams**. 1989. General cytology of fission yeasts. p. 273-330. In A. Nasim, P. Young, and B.F. Johnson. (ed.), Molecular biology of the fission yeast. Academic Press, Inc., San Diego, California.

**Roise, D. and G. Schatz**. 1988. Mitochondrial presequences. J.Biol.Chem. **263**:4509-4511.

**Rossi, R., A. Villa, C. Negri, I. Scovassi, G. Ciarrocchi, G. Biamonti, and A. Montecucco**. 1999. The replication factory targeting sequence/PCNA-binding site is required in G(1) to control the phosphorylation status of DNA ligase I. EMBO J. **18**:5745-5754.

**Sambrook, J., E.F. Fritsch, and T. Maniatis**. 1989. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

**Savini, E., G. Biamonti, G. Ciarrocchi, and A. Montecucco**. 1994. Cloning and sequence analysis of a cDNA coding for the murine DNA ligase I enzyme. Gene **144**:253-257.

**Sawin, K.E. and P. Nurse**. 1996. Identification of fission yeast nuclear markers using random polypeptide fusions with green fluorescent protein. Proc.Natl.Acad.Sci.U.S.A. **93**:15146-15151.

**Sawin, K.E.** 1999. GFP fusion proteins as probes for cytology in fission yeast. Methods Cell Biol. **58**:123-138.

**Schar, P., G. Herrmann, G. Daly, and T. Lindahl**. 1997. A newly identified DNA ligase of *Saccharomyces cerevisiae* involved in RAD52-independent repair of DNA double-strand breaks. Genes Dev. **11**:1912-1924.

**Sekiguchi, J. and S. Shuman**. 1997. Domain structure of vaccinia DNA ligase. Nucleic.Acids.Res. **25**:727-734.

**Shcherbakova, P.V. and Y.I. Pavlov**. 1996. 3'-->5' exonucleases of DNA polymerases epsilon and delta correct base analog induced DNA replication errors on opposite DNA strands in *Saccharomyces cerevisiae*. Genetics **142**:717-726.

**Shikata, K., S. Ohta, K. Yamada, C. Obuse, H. Yoshikawa, and T. Tsurimoto**. 2001. The human homologue of fission Yeast cdc27, p66, is a component of active human DNA polymerase delta. J.Biochem.(Tokyo.) **129**:699-708.

**Shuman, S. and B. Schwer**. 1995. RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases. Mol.Microbiol. **17**:405-410.

**Sibanda, B.L., S.E. Critchlow, J. Begun, X.Y. Pei, S.P. Jackson, T.L. Blundell, and L. Pellegrini**. 2001. Crystal structure of an Xrcc4-DNA ligase IV complex. Nat.Struct.Biol. **8**:1015-1019

**Singleton, M.R., K. Hakansson, D.J. Timson, and D.B. Wigley**. 1999. Structure of the adenylation domain of an NAD+-dependent DNA ligase. Structure.Fold.Des. **7**:35-42.

**Sipiczki, M.** 2000. Where does fission yeast sit on the tree of life? Genome Biol. **1**:REVIEWS1011

**Soderhall, S. and T. Lindahl**. 1975. Mammalian DNA ligases. Serological evidence for two separate enzymes. J.Biol.Chem. **250**:8438-8444.

**Soderhall, S.** 1975. Properties of a DNA-adenylate complex formed in the reaction between mammalian DNA ligase I and DNA containing single-strand breaks. Eur.J.Biochem. **51**:UNKNOWN

**Soderhall, S. and T. Lindahl**. 1976. DNA ligases of eukaryotes. FEBS Lett. **67**:1-8.

**Sommers, C.H., E.J. Miller, B. Dujon, S. Prakash, and L. Prakash**. 1995. Conditional lethality of null mutations in RTH1 that encodes the yeast counterpart of a mammalian 5'- to 3'-exonuclease required for lagging strand DNA synthesis in reconstituted systems. J.Biol.Chem. **270**:4193-4196.

**Sriskanda, V. and S. Shuman**. 1998a. Chlorella virus DNA ligase: nick recognition and mutational analysis. Nucleic.Acids.Res. **26**:525-531.

**Sriskanda, V. and S. Shuman**. 1998b. Mutational analysis of Chlorella virus DNA ligase: catalytic roles of domain I and motif VI. Nucleic.Acids.Res. **26**:4618-4625.

**Sriskanda, V., B. Schwer, C.K. Ho, and S. Shuman**. 1999. Mutational analysis of Escherichia coli DNA ligase identifies amino acids required for nick-ligation *in vitro* and for *in vivo* complementation of the growth of yeast cells deleted for CDC9 and LIG4. Nucleic.Acids.Res. **27**:3953-3963.

**Sriskanda, V., R.W. Moyer, and S. Shuman**. 2001. NAD+-dependent DNA ligase encoded by a eukaryotic virus. J.Biol.Chem. **276**:36100-36109.

**Stucki, M., Z.O. Jonsson, and U. Hubscher**. 2001. In eukaryotic flap endonuclease 1, the C terminus is essential for substrate binding. J.Biol.Chem. **276**:7843-7849.

**Studier, F.W. and B.A. Moffatt**. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J.Mol.Biol. **189**:113-130.

**Subramanya, H.S., A.J. Doherty, S.R. Ashford, and D.B. Wigley**. 1996. Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. Cell **85**:607-615.

**Sugimoto, K., T. Shimomura, K. Hashimoto, H. Araki, A. Sugino, and K. Matsumoto**. 1996. Rfc5, a small subunit of replication factor C complex, couples DNA replication and mitosis in budding yeast. Proc.Natl.Acad.Sci.U.S.A. **93**:7048-7052.

**Tanaka, T., D. Knapp, and K. Nasmyth**. 1997. Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs. Cell **90**:649-660.

**Taylor, R.M., B. Wickstead, S. Cronin, and K.W. Caldecott**. 1998a. Role of a BRCT domain in the interaction of DNA ligase III-alpha with the DNA repair protein XRCC1. Curr.Biol. **8**:877-880.

**Taylor, R.M., M.J. Hamer, J. Rosamond, and C.M. Bray**. 1998b. Molecular cloning and functional analysis of the Arabidopsis thaliana DNA ligase I homologue. Plant J. **14**:75-81.

**Teo, S.H. and S.P. Jackson**. 1997. Identification of *Saccharomyces cerevisiae* DNA ligase IV: involvement in DNA double-strand break repair. EMBO J. **16**:4788-4795.

**Teraoka, H. and K. Tsukada**. 1982. Eukaryotic DNA ligase. Purification and properties of the enzyme from bovine thymus, and immunochemical studies of the enzyme from animal tissues. J.Biol.Chem. **257**:4758-4763.

**Teraoka, H., T. Sumikawa, and K. Tsukada**. 1986. Purification of DNA ligase II from calf thymus and preparation of rabbit antibody against calf thymus DNA ligase II. J.Biol.Chem. **261**:6888-6892.

**Timson, D.J., M.R. Singleton, and D.B. Wigley**. 2000. DNA ligases in the repair and replication of DNA. Mutat.Res. **460**:301-318.

**Tom, S., L.A. Henricksen, M.S. Park, and R.A. Bambara**. 2001. DNA ligase I and proliferating cell nuclear antigen form a functional complex. J.Biol.Chem. **276**:24817-24825.

**Tomkinson, A.E., L. Chen, Z. Dong, J.B. Leppard, D.S. Levin, Z.B. Mackey, and T.A. Motycka**. 1902. Completion of base excision repair by mammalian DNA ligases. Prog.Nucleic.Acid.Res.Mol.Biol. **68**:151-164.

**Tomkinson, A.E., D.D. Lasko, G. Daly, and T. Lindahl**. 1990. Mammalian DNA ligases. Catalytic domain and size of DNA ligase I. J.Biol.Chem. **265**:12611-12617.

**Tomkinson, A.E., E. Roberts, G. Daly, N.F. Totty, and T. Lindahl**. 1991. Three distinct DNA ligases in mammalian cells. J.Biol.Chem. **266**:21728-21735.

**Tomkinson, A.E., N.F. Totty, M. Ginsburg, and T. Lindahl** . 1991. Location of the active site for enzyme-adenylate formation in DNA ligases. Proc.Natl.Acad.Sci.U.S.A. **88**:400-404.

**Tomkinson, A.E., N.J. Tappe, and E.C. Friedberg**. 1992. DNA ligase I from *Saccharomyces cerevisiae*: physical and biochemical characterization of the CDC9 gene product. Biochemistry **31**:11762-11771.

**Tomkinson, A.E. and D.S. Levin**. 1997. Mammalian DNA ligases. Bioessays **19**:893-901.

**Tomkinson, A.E. and Z.B. Mackey**. 1998. Structure and function of mammalian DNA ligases. Mutat.Res. **407**:1-9.

**Tournier, S., D. Leroy, F. Goubin, B. Ducommun, and J.S. Hyams**. 1996. Heterologous expression of the human cyclin-dependent kinase inhibitor p21Cip1 in the fission yeast,

*Schizosaccharomyces pombe* reveals a role for PCNA in the chk1+ cell cycle checkpoint pathway. Mol.Biol.Cell 7:651-662.

**Tratner, I., K. Piard, M. Grenon, M. Perderiset, and G. Baldacci**. 1997. PCNA and DNA polymerase delta catalytic subunit from *Schizosaccharomyces pombe* do not interact directly. Biochem.Biophys.Res.Commun. **231**:321-328.

**Tsuchimoto, D., Y. Sakai, K. Sakumi, K. Nishioka, M. Sasaki, T. Fujiwara, and Y. Nakabeppu**. 2001. Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen. Nucleic.Acids.Res. **29**:2349-2360.

**Tsurimoto, T. and B. Stillman**. 1989. Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication *in vitro*. Mol.Cell Biol. **9**:609-619.

**Tsurimoto, T. and B. Stillman**. 1990. Functions of replication factor C and proliferating-cell nuclear antigen: functional similarity of DNA polymerase accessory proteins from human cells and bacteriophage T4. Proc.Natl.Acad.Sci.U.S.A. **87**:1023-1027.

**Tsurimoto, T. and B. Stillman**. 1991a. Replication factors required for SV40 DNA replication *in vitro*. I. DNA structure-specific recognition of a primer-template junction by eukaryotic DNA polymerases and their accessory proteins. J.Biol.Chem. **266**:1950-1960.

**Tsurimoto, T. and B. Stillman**. 1991b. Replication factors required for SV40 DNA replication *in vitro*. II. Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand synthesis. J.Biol.Chem. **266**:1961-1968.

**Tsurimoto, T.** 1999. PCNA binding proteins. Front.Biosci. **4**:D849-D858

**Turchi, J.J., L. Huang, R.S. Murante, Y. Kim, and R.A. Bambara**. 1994. Enzymatic completion of mammalian lagging-strand DNA replication. Proc.Natl.Acad.Sci.U.S.A. **91**:9803-9807.

**Uhlmann, F., J. Cai, H. Flores-Rozas, F.B. Dean, J. Finkelstein, M. O'Donnell, and J. Hurwitz**. 1996. *In vitro* reconstitution of human replication factor C from its five subunits. Proc.Natl.Acad.Sci.U.S.A. **93**:6521-6526.

**Uhlmann, F., J. Cai, E. Gibbs, M. O'Donnell, and J. Hurwitz**. 1997. Deletion analysis of the large subunit p140 in human replication factor C reveals regions required for complex formation and replication activities. J.Biol.Chem. **272**:10058-10064.

**Uhlmann, F., E. Gibbs, J. Cai, M. O'Donnell, and J. Hurwitz**. 1997. Identification of regions within the four small subunits of human replication factor C required for complex formation and DNA replication. J.Biol.Chem. **272**:10065-10071.

**Vashee, S., P. Simancek, M.D. Challberg, and T.J. Kelly**. 2001. Assembly of the human origin recognition complex. J.Biol.Chem. **276**:26666-26673.

**von Heijne, G.** 1986. Mitochondrial targeting sequences may form amphiphilic helices. EMBO J. **5**:1335-1342.

**Waga, S. and B. Stillman**. 1994. Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication *in vitro*. Nature **369**:207-212.

**Waga, S., G. Bauer, and B. Stillman**. 1994. Reconstitution of complete SV40 DNA replication with purified replication factors. J.Biol.Chem. **269**:10923-10934.

**Waga, S. and B. Stillman**. 1998. The DNA replication fork in eukaryotic cells. Annu.Rev.Biochem. **67**:721-751.

**Wang, T.S.** 1991. Eukaryotic DNA polymerases. Annu.Rev.Biochem. **60**:513-552.

**Warbrick, E., D.P. Lane, D.M. Glover, and L.S. Cox**. 1995. A small peptide inhibitor of DNA replication defines the site of interaction between the cyclin-dependent kinase inhibitor p21WAF1 and proliferating cell nuclear antigen. Curr.Biol. **5**:275-282.

**Warbrick, E., D.P. Lane, D.M. Glover, and L.S. Cox**. 1997. Homologous regions of Fen1 and p21Cip1 compete for binding to the same site on PCNA: a potential mechanism to co-ordinate DNA replication and repair. Oncogene **14**:2313-2321.

**Warbrick, E.** 1998. PCNA binding through a conserved motif. Bioessays **20**:195-199.

**Warbrick, E.** 2000. The puzzle of PCNA's many partners. Bioessays **22**:997-1006.

**Waseem, N.H., K. Labib, P. Nurse, and D.P. Lane**. 1992. Isolation and analysis of the fission yeast gene encoding polymerase delta accessory protein PCNA. EMBO J. **11**:5111-5120.

**Wei, Y.F., P. Robins, K. Carter, K. Caldecott, D.J. Pappin, G.L. Yu, R.P. Wang, B.K. Shell, R.A. Nash, and P. Schar**. 1995. Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination. Mol.Cell Biol. **15**:3206-3216.

**White, J.H., D.G. Barker, P. Nurse, and L.H. Johnston**. 1986. Periodic transcription as a means of regulating gene expression during the cell cycle: contrasting modes of expression of DNA ligase genes in budding and fission yeast. EMBO J. **5**:1705-1709.

**Willer, M., M. Rainey, T. Pullen, and C.J. Stirling**. 1999. The yeast CDC9 gene encodes both a nuclear and a mitochondrial form of DNA ligase I. Curr.Biol. **9**:1085-1094.

**Wilson, T.E., U. Grawunder, and M.R. Lieber**. 1997. Yeast DNA ligase IV mediates non-homologous DNA end joining. Nature **388**:495-498.

**Wood, R.D.** 1996. DNA repair in eukaryotes. Annu.Rev.Biochem. **65**:135-167.

**Xiong, Y., G.J. Hannon, H. Zhang, D. Casso, R. Kobayashi, and D. Beach**. 1993. p21 is a universal inhibitor of cyclin kinases. Nature **366**:701-704.

**Yaffe, M.P., D. Harata, F. Verde, M. Eddison, T. Toda, and P. Nurse**. 1996. Microtubules mediate mitochondrial distribution in fission yeast. Proc.Natl.Acad.Sci.U.S.A. **93**:11664-11668.

**Yanisch-Perron, C., J. Vieira, and J. Messing**. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103-119.

**Yasuhira, S. and A. Yasui**. 2000. Alternative excision repair pathway of UV-damaged DNA in *Schizosaccharomyces pombe* operates both in nucleus and in mitochondria. J.Biol.Chem. **275**:11824-11828.

**Yuzhakov, A., Z. Kelman, J. Hurwitz, and M. O'Donnell**. 1999. Multiple competition reactions for RPA order the assembly of the DNA polymerase delta holoenzyme. EMBO J. **18**:6189-6199.

**Zhang, P., J.Y. Mo, A. Perez, A. Leon, L. Liu, N. Mazloum, H. Xu, and M.Y. Lee**. 1999. Direct interaction of proliferating cell nuclear antigen with the p125 catalytic subunit of mammalian DNA polymerase delta. J.Biol.Chem. **274**:26647-26653.

**Zhang, S.J., X.R. Zeng, P. Zhang, N.L. Toomey, R.Y. Chuang, L.S. Chang, and M.Y. Lee**. 1995. A conserved region in the amino terminus of DNA polymerase delta is involved in proliferating cell nuclear antigen binding. J.Biol.Chem. **270**:7988-7992.

**Zhou, J.Q., H. He, C.K. Tan, K.M. Downey, and A.G. So**. 1997. The small subunit is required for functional interaction of DNA polymerase delta with the proliferating cell nuclear antigen. Nucleic.Acids.Res. **25**:1094-1099.

# Appendix 1: Nucleotide sequence of *cdc17*

```
         10         20         30         40         50         60         70
         |          |          |          |          |          |          |
AAGCTTCCAATGACGGAGAAAGAAATTGCCATTTGAAGAGCCAGAGCTGGTAGATTGAAAACGTACAGAGGAGCGCTT

 80         90         100        110        120        130        140        150
 |          |          |          |          |          |          |          |
TAAACCCTGAGCGACACCGCGAAACAAAGATCGTTGATACATTTTAGTATTTTTAATAGAATTCGGCTTAAATAAAAA

   160        170        180        190        200        210        220        230
   |          |          |          |          |          |          |          |
ATTTTCAAAACGAAATTGAAAAAAATATGAATGATAAAATAATAAAATGAAGTTTTATAATAGTTTATGCTTAATATA

     240        250        260        270        280        290        300        310
     |          |          |          |          |          |          |          |
TTTAAACACCAGTTTTTTTCCAAACACAACAAACCCAAGAAAAAATAAACAGAACCAAAAAATAAAAATATGGATTTA

       320        330        340        350        360        370        380        390
       |          |          |          |          |          |          |          |
AGGATCAGAACAGAATCGTCAATCCGTTTCTTTATCCACCTAGAGACAAGAAGTGTTAATTTATATAAAATGGAGAAC

         400        410        420        430        440        450        460
         |          |          |          |          |          |          |
TTGGTTCTAATAAGATTTCAACGAAAACAAAGATTTCAGAACCGAAAAAATTTCTCCAAGTGCTAGCAGAGGTATATT

 470        480        490        500        510        520        530        540
 |          |          |          |          |          |          |          |
CTCTGAGAAAGAACAAAAATCAAAAATTGAATGTTTATATTGAATGTTACCTGTAACAGGTAACATGTTTTTTTAGTA

   550        560        570        580        590        600        610        620
   |          |          |          |          |          |          |          |
ATGCCGGGCTGTCGAAATTTGACTATGAGTTTTGTAAATAAAGCGATGGTTATAAACAATTTTATGATAGTTTTTGCT

     630        640        650        660        670        680        690        700
     |          |          |          |          |          |          |          |
TTCGAATTAACATTTCATAAAAAATTTGAAAAAGAAAAATTTTAAAAATTTAAGGACAAAATCCACTTGAAAAACAAG

       710        720        730        740        750        760        770        780
       |          |          |          |          |          |          |          |
TGCGAACAGCTTTTTGCAAAAAAACTACAAACCACTACACATTAGCGAATAAGCCTATTAGGAATTTATTTTAATTTT

         790        800        810        820        830        840        850
         |          |          |          |          |          |          |
TGTTGATTTAAATTTCTTTTTTTTTTAAATGGTTAGAATCTAGTTGATAAACTTCTAGCAATGTATTTCCAAATTGTT

 860        870        880        890        900        910        920        930
 |          |          |          |          |          |          |          |
TCTATTACGTTCGTTGATACAGGCAATTGTTCATTTCTGGTATGACTAAAGCATTTCTATACAATTCAACTTCTCATT

   940        950        960        970        980        990        1000       1010
   |          |          |          |          |          |          |          |
TTTTCTAGCTGTATTTATGATGTACAGAGCGTTGGGAGAGAGTTGATTAAGTCAATAGTATTTATGAATGAGACAAGG

     1020       1030       1040       1050       1060       1070       1080       1090
     |          |          |          |          |          |          |          |
AACACTCATACTATATCAATGATAAATAGCTTTTATATTATCCTGTTATAATCATAAAAATTAATTTCATCTATATAG

       1100       1110       1120       1130       1140       1150       1160       1170
       |          |          |          |          |          |          |          |
ACTTAGAAAAAAATGGTTTCGTTTTCACTGTCCTTAAACTACAAATATCGTACACGTGTTACTTGTTCTTCAGGTGTA

         1180       1190       1200       1210       1220       1230       1240
         |          |          |          |          |          |          |
CCCTCACTATTGTCAAATCTTATGAAATTCTTTACTCGTTGACTTGATTGTTATGCGAACAGTATTTTCGCAAATTCC
                                                    M  R  T  V  F  S  Q  I  P

 1250       1260       1270       1280       1290       1300       1310       1320
 |          |          |          |          |          |          |          |
TCGTTTTAAACAAGTCAATCAGTACATACGGATGTCGACAAGGCAGTCAGACATTAGGTTTGTTAGATAGAAGACAGA
 R  F  K  Q  V  N  Q  Y  I  R  M  S  T  R  Q  S  D  I
```

```
   1330       1340       1350       1360       1370       1380       1390       1400
   |          |          |          |          |          |          |          |
CGTTGACTAACCCATATAGCAACTTCTTCATTTCTTCTGCTTCGCACAAAAGCGAGCACGTGGAGGTTTCTCAATCGT
                S  N  F  F  I  S  S  A  S  H  K  S  E  H  V  E  V  S  Q  S  S

     1410       1420       1430       1440       1450       1460       1470       1480
     |          |          |          |          |          |          |          |
CTTCTGATTCCAAAAATGTTGATGGGAGAAGCACAAGCGAAAAGAGAAAGGTCGAATCAGTGAAATTAGTCGACGAGT
  S  D  S  K  N  V  D  G  R  S  T  S  E  K  R  K  V  E  S  V  K  L  V  D  E  S

       1490       1500       1510       1520       1530       1540       1550       1560
       |          |          |          |          |          |          |          |
CGAAACACAACAACCACGACGACACTGGAACTCAAAATGTTGAGCGTGAAAATAATATTGTGTCTGAGGCAAAAAAGC
  K  H  N  N  H  D  D  T  G  T  Q  N  V  E  R  E  N  N  I  V  S  E  A  K  K  Q

         1570       1580       1590       1600       1610       1620       1630
         |          |          |          |          |          |          |
AAAAGACTTTAGGTTCATCCTCCTCTTCTTCAGATGCCGTATCTAGTAACAACGATTCTGGGGCATCTACTCCAATAC
  K  T  L  G  S  S  S  S  S  S  D  A  V  S  S  N  N  D  S  G  A  S  T  P  I  P

 1640       1650       1660       1670       1680       1690       1700       1710
 |          |          |          |          |          |          |          |
CTTTGCCAATCAAAGAACCACCGTTAGAGTCCAATGCTCGCAATGATAAGTTAAAAGGTCATGCAACATTTGCTGAAA
  L  P  I  K  E  P  P  L  E  S  N  A  R  N  D  K  L  K  G  H  A  T  F  A  E  M

   1720       1730       1740       1750       1760       1770       1780       1790
   |          |          |          |          |          |          |          |
TGGTAAAAGCATTCACAAAAATAGAGAATACTAGCAAACGGTTGGAAATCATTGATATAATGGGGACTTACTTTTTTG
  V  K  A  F  T  K  I  E  N  T  S  K  R  L  E  I  I  D  I  M  G  T  Y  F  F  G

     1800       1810       1820       1830       1840       1850       1860       1870
     |          |          |          |          |          |          |          |
GAATTCTTCGCGACCATCCCAGCGATTTACTGGCCTGTGTGTATTTAAGTATCAATAAGGTATGAACGTTTGGTTTTA
  I  L  R  D  H  P  S  D  L  L  A  C  V  Y  L  S  I  N  K

       1880       1890       1900       1910       1920       1930       1940       1950
       |          |          |          |          |          |          |          |
AAGGATTTCTCCTTCTCAGTACTAACGTCTTTTTTTACATTAGTTGGGCCCAGATTACTCGGGACTAGAGCTAGGAAT
                                          L  G  P  D  Y  S  G  L  E  L  G  I

         1960       1970       1980       1990       2000       2010       2020
         |          |          |          |          |          |          |
TGGCGAAAGCATTATAATGAAGGCTATTGGAGAGTCTACAGGGCAAACGCTTCAACAGATTAAACTATCGTTTCACAA
 G  E  S  I  I  M  K  A  I  G  E  S  T  G  Q  T  L  Q  Q  I  K  L  S  F  H  K

 2030       2040       2050       2060       2070       2080       2090       2100
 |          |          |          |          |          |          |          |
AGTTGGTGATCTTGGTCTTGTAGCTCAAACATCCCGTCAAAATCAACCCACTATGTTTAAACCTGCAGCGCTGACAAT
 V  G  D  L  G  L  V  A  Q  T  S  R  Q  N  Q  P  T  M  F  K  P  A  A  L  T  I

   2110       2120       2130       2140       2150       2160       2170       2180
   |          |          |          |          |          |          |          |
TCCATTTTTATTTGACTCATTAAAAAAAATTGCTCAAATGAGCGGAAACCAATCCCAAAATAGAAAAATTGGAGTGAT
 P  F  L  F  D  S  L  K  K  I  A  Q  M  S  G  N  Q  S  Q  N  R  K  I  G  V  I

     2190       2200       2210       2220       2230       2240       2250       2260
     |          |          |          |          |          |          |          |
CAAGCGGCTATTGAGTTCCTGTGAAGGTGCAGAACCTAAATATTTGATTCGAGCTTTGGAAGGCAAGCTACGCCTCCA
 K  R  L  L  S  S  C  E  G  A  E  P  K  Y  L  I  R  A  L  E  G  K  L  R  L  Q

       2270       2280       2290       2300       2310       2320       2330       2340
       |          |          |          |          |          |          |          |
ACTTGCCGAAAAAACCATTTTAGTAGCATTAGCAAATGCCACCGCTCAATATCATGCTGATAAGAATGGAGAAAAACT
 L  A  E  K  T  I  L  V  A  L  A  N  A  T  A  Q  Y  H  A  D  K  N  G  E  K  L

         2350       2360       2370       2380       2390       2400       2410
         |          |          |          |          |          |          |
TTCGCAACAAGACAGGATTGAAGGAGAGCAGATTCTTCGAGATGTTTACTGTCAACTGCCTTCATACGACCTTATCGT
 S  Q  Q  D  R  I  E  G  E  Q  I  L  R  D  V  Y  C  Q  L  P  S  Y  D  L  I  V

 2420       2430       2440       2450       2460       2470       2480       2490
 |          |          |          |          |          |          |          |
TCCTCACTTGATTGAGCATGGACTTGGTACTTTGCGTGAAACTTGTAAATTAACACCTGGAATCCCAACAAAACCCAT
```

```
 P  H  L  I  E  H  G  L  G  T  L  R  E  T  C  K  L  T  P  G  I  P  T  K  P  M

   2500       2510       2520       2530       2540       2550       2560       2570
   |          |          |          |          |          |          |          |
GCTGGCTAAGCCGACGAAACAAATTTCAGAAGTTTTGAATACGTTTGATCAAGCCGCATTTACATGTGAATACAAATA
 L  A  K  P  T  K  Q  I  S  E  V  L  N  T  F  D  Q  A  A  F  T  C  E  Y  K  Y

     2580       2590       2600       2610       2620       2630       2640       2650
     |          |          |          |          |          |          |          |
CGATGGCGAGCGCGCCCAAGTTCATTTTACCGAAGATGGGAAGTTTTATGTATTTTCTAGGAATTCTGAAAACATGTC
 D  G  E  R  A  Q  V  H  F  T  E  D  G  K  F  Y  V  F  S  R  N  S  E  N  M  S

       2660       2670       2680       2690       2700       2710       2720       2730
       |          |          |          |          |          |          |          |
TGTTCGCTATCCGGACATCTCTGTTTCTGTATCGAAGTGGAAAAAACCAGATGCTCGTTCTTTTATTCTCGATTGCGA
 V  R  Y  P  D  I  S  V  S  V  S  K  W  K  K  P  D  A  R  S  F  I  L  D  C  E

         2740       2750       2760       2770       2780       2790       2800
         |          |          |          |          |          |          |
AGCTGTCGGTTGGGACCGTGATGAAAACAAAATTTTACCTTTTCAAAAACTTGCAACAAGAAAGAGGAAAGATGTGAA
 A  V  G  W  D  R  D  E  N  K  I  L  P  F  Q  K  L  A  T  R  K  R  K  D  V  K

 2810       2820       2830       2840       2850       2860       2870       2880
 |          |          |          |          |          |          |          |
AATTGGGGACATTAAGGTCCGAGCATGCCTTTTCGCTTTTGACATTTTATACCTAAATGGGCAACCTCTTTTGGAAAC
 I  G  D  I  K  V  R  A  C  L  F  A  F  D  I  L  Y  L  N  G  Q  P  L  L  E  T

   2890       2900       2910       2920       2930       2940       2950       2960
   |          |          |          |          |          |          |          |
CCCACTCAACGAGCGTCGAAAACTACTTTACTCAATGTTTCAACCAAGTACTGGAGACTTTACTTTTGCTAAACATAG
 P  L  N  E  R  R  K  L  L  Y  S  M  F  Q  P  S  T  G  D  F  T  F  A  K  H  S

     2970       2980       2990       3000       3010       3020       3030       3040
     |          |          |          |          |          |          |          |
CGATCAAAAGTCTATTGAATCCATTGAAGAATTCCTTGAAGAATCTGTCAAAGATTCTTGTGAGGGTTTAATGGTCAA
 D  Q  K  S  I  E  S  I  E  E  F  L  E  E  S  V  K  D  S  C  E  G  L  M  V  K

       3050       3060       3070       3080       3090       3100       3110       3120
       |          |          |          |          |          |          |          |
AATGCTTGAAGGCCCGGATTCTCATTATGAACCGTCAAAACGTTCCAGGCATTGGCTAAAAGTTAAAAAAGACTATCT
 M  L  E  G  P  D  S  H  Y  E  P  S  K  R  S  R  H  W  L  K  V  K  K  D  Y  L

         3130       3140       3150       3160       3170       3180       3190
         |          |          |          |          |          |          |
TTCTGGTGTTGGAGATTCTCTGGATTTGATTGTAATTGGCGCATACTACGGAAAGGGTAAACGTACTTCTGTTTATGG
 S  G  V  G  D  S  L  D  L  I  V  I  G  A  Y  Y  G  K  G  K  R  T  S  V  Y  G

 3200       3210       3220       3230       3240       3250       3260       3270
 |          |          |          |          |          |          |          |
TGCCTTTCTTTTGGGCTGCTACGATCCAGATACCGAAACAGTTCAATCGATTTGCAAACTCGGGACTGGTTTTTCAGA
 A  F  L  L  G  C  Y  D  P  D  T  E  T  V  Q  S  I  C  K  L  G  T  G  F  S  E

   3280       3290       3300       3310       3320       3330       3340       3350
   |          |          |          |          |          |          |          |
GGAGCATTTAGAAACTTTTTACAATCAACTTAAAGACATTGTTATATCTAAAAAGAAAGATTTTTATGCTCATAGCGA
 E  H  L  E  T  F  Y  N  Q  L  K  D  I  V  I  S  K  K  K  D  F  Y  A  H  S  D

     3360       3370       3380       3390       3400       3410       3420       3430
     |          |          |          |          |          |          |          |
TGTACCTGCGCATCAGCCTGACGTCTGGTTTGAGCCTAAGTACTTGTGGGAAGTCTTAGCTGCCGATTTATCTTTGTC
 V  P  A  H  Q  P  D  V  W  F  E  P  K  Y  L  W  E  V  L  A  A  D  L  S  L  S

       3440       3450       3460       3470       3480       3490       3500       3510
       |          |          |          |          |          |          |          |
TCCGGTTTACAAAGCTGCAATTGGATATGTTCAAGAGGACAAAGGAATCAGTTTACGTTTTCCAAGATTCATTCGTAT
 P  V  Y  K  A  A  I  G  Y  V  Q  E  D  K  G  I  S  L  R  F  P  R  F  I  R  I

         3520       3530       3540       3550       3560       3570       3580
         |          |          |          |          |          |          |
TCGGGAAGACAAGAGCTGGGAGGATGCTACAACAAGTGAACAGGTTTCAGAATTCTATCGTTCGCAGGTGGCTTACAG
 R  E  D  K  S  W  E  D  A  T  T  S  E  Q  V  S  E  F  Y  R  S  Q  V  A  Y  S

 3590       3600       3610       3620       3630       3640       3650       3660
 |          |          |          |          |          |          |          |
```

```
TCAAAAGGAAAAAGAAGGGTCCCCAGCTGCCGAAGATTATTAAGTTTGTAGCTGAGAGTTACTGCTGAGAGCTACTGT
 Q  K  E  K  E  G  S  P  A  A  E  D  Y  ¥

   3670       3680       3690       3700       3710       3720       3730       3740
   |          |          |          |          |          |          |          |
CATAATAAAGGGTAATGTTCGGCTAAAATTTGCCCAAAGAGGCTAATATGAATTATGTTAAAGAGACAGACAAATTGA

     3750       3760       3770       3780       3790       3800       3810       3820
     |          |          |          |          |          |          |          |
GTTTATGCTCTAGCTGATGGCTATTTCTCTAGCTTACTACTATTTATTTTTTTCTAATTGTAATCCCCTGTGGCTAAA

       3830       3840       3850       3860       3870       3880       3890       3900
       |          |          |          |          |          |          |          |
AAAAAATTAATAATAAACAACATTGAATTTCCAAATTATAACATTAGATTGAAAATTTGAATCTTATTTATCCCTTTT

         3910       3920       3930       3940       3950       3960       3970
         |          |          |          |          |          |          |
TATTATCCTTTTTATATGATTTTATACATTGTATTACCATTTTACGTTCGTACCAAGCCTTACTCTATTTTGCCGTCA

 3980       3990       4000       4010       4020       4030       4040       4050
 |          |          |          |          |          |          |          |
TATTCCTCCTTACAGTGGATTACGAATGCTTGGTGTTCATTACCGTCCATATACATTTTATGTCACTGTAGCAAACAC

   4060       4070       4080       4090       4100       4110       4120       4130
   |          |          |          |          |          |          |          |
CAAGAAGAAATTACAATATTGGCTTTAAAGCCATTAATTGCTATGAATAATTCAGAAGAATGGCCATTGAAACATGTG

     4140       4150       4160       4170
     |          |          |          |
CAACCGCCCAAATTGTCATTTCGTCATTTGCTATCGAGATCT
```

## Appendix 2: Protein sequence of Cdc17

```
1           11          21          31          41          51          61
MRTVFSQIPR FKQVNQYIRM STRQSDISNF FISSASHKSE HVEVSQSSSD SKNVDGRSTS EKRKVESVKL

71          81          91          101         111         121         131
VDESKHNNHD DTGTQNVERE NNIVSEAKKQ KTLGSSSSSS DAVSSNNDSG ASTPIPLPIK EPPLESNARN

141         151         161         171         181         191         201
DKLKGHATFA EMVKAFTKIE NTSKRLEIID IMGTYFFGIL RDHPSDLLAC VYLSINKLGP DYSGLELGIG

211         221         231         241         251         261         271
ESIIMKAIGE STGQTLQQIK LSFHKVGDLG LVAQTSRQNQ PTMFKPAALT IPFLFDSLKK IAQMSGNQSQ

281         291         301         311         321         331         341
NRKIGVIKRL LSSCEGAEPK YLIRALEGKL RLQLAEKTIL VALANATAQY HADKNGEKLS QQDRIEGEQI

351         361         371         381         391         401         411
LRDVYCQLPS YDLIVPHLIE HGLGTLRETC KLTPGIPTKP MLAKPTKQIS EVLNTFDQAA FTCEYKYDGE

421         431         441         451         461         471         481
RAQVHFTEDG KFYVFSRNSE NMSVRYPDIS VSVSKWKKPD ARSFILDCEA VGWDRDENKI LPFQKLATRK

491         501         511         521         531         541         551
RKDVKIGDIK VRACLFAFDI LYLNGQPLLE TPLNERRKLL YSMFQPSTGD FTFAKHSDQK SIESIEEFLE

561         571         581         591         601         611         621
ESVKDSCEGL MVKMLEGPDS HYEPSKRSRH WLKVKKDYLS GVGDSLDLIV IGAYYGKGKR TSVYGAFLLG

631         641         651         661         671         681         691
CYDPDTETVQ SICKLGTGFS EEHLETFYNQ LKDIVISKKK DFYAHSDVPA HQPDVWFEPK YLWEVLAADL

701         711         721         731         741         751         761
SLSPVYKAAI GYVQEDKGIS LRFPRFIRIR EDKSWEDATT SEQVSEFYRS QVAYSQKEKE GSPAAEDY
```

## Appendix 3: Data for growth curves of *cdc17$^{ts}$* strains described in section 2.2.2

***cdc17-K42***

| time / h | OD 25°C | OD 36°C | cell no. 25°C [x10E6] | cell no. 36°C [x10E6] |
|---|---|---|---|---|
| 0 | 0,055 | 0,045 | 0,5403 | 0,4808 |
| 1 | 0,062 | 0,055 | 0,5438 | 0,4549 |
| 2 | 0,074 | 0,071 | 0,5002 | 0,3589 |
| 3 | 0,079 | 0,103 | 0,4655 | 0,4406 |
| 4 | 0,099 | 0,124 | 0,8238 | 0,6985 |
| 5 | 0,127 | 0,166 | 0,9145 | 0,7445 |
| 6 | 0,147 | 0,193 | 1,35 | 0,7345 |
| 7 | 0,2 | 0,237 | 1,756 | 0,8376 |
| 8 | 0,263 | 0,284 | 2,233 | 0,9004 |
| 9 | 0,342 | 0,31 | 3,334 | 1,062 |
| 10 | 0,434 | 0,374 | 3,912 | 1,065 |

***cdc17-L16***

| time / h | OD 25°C | OD 36°C | cell no. 25°C [x10E6] | cell no. 36°C [x10E6] |
|---|---|---|---|---|
| 0 | 0,050 | 0,040 | 0,5366 | 0,4670 |
| 1 | 0,071 | 0,063 | 0,6654 | 0,4381 |
| 2 | 0,072 | 0,071 | 0,4745 | 0,3045 |
| 3 | 0,080 | 0,093 | 0,4536 | 0,4637 |
| 4 | 0,094 | 0,120 | 0,7045 | 0,5541 |
| 5 | 0,118 | 0,151 | 0,8905 | 0,7327 |
| 6 | 0,143 | 0,190 | 1,1670 | 0,7667 |
| 7 | 0,181 | 0,243 | 1,4710 | 0,8708 |
| 8 | 0,235 | 0,307 | 1,7200 | 1,0379 |
| 9 | 0,288 | 0,349 | 2,3780 | 1,2160 |
| 10 | 0,356 | 0,406 | 2,9590 | 1,2250 |

***cdc17-M75***

| time / h | OD 25°C | OD 36°C | cell no. 25°C [x10E6] | cell no. 36°C [x10E6] |
|---|---|---|---|---|
| 0 | 0,054 | 0,053 | 0,5262 | 0,5652 |
| 1 | 0,063 | 0,079 | 0,5823 | 0,5908 |
| 2 | 0,075 | 0,089 | 0,5317 | 0,4763 |
| 3 | 0,087 | 0,118 | 0,5005 | 0,5195 |
| 4 | 0,109 | 0,162 | 0,8473 | 0,8696 |
| 5 | 0,145 | 0,208 | 1,1020 | 1,0330 |
| 6 | 0,165 | 0,270 | 1,3660 | 1,3890 |
| 7 | 0,206 | 0,346 | 1,7930 | 1,6580 |
| 8 | 0,274 | 0,450 | 2,4670 | 1,9190 |
| 9 | 0,331 | 0,543 | 3,4050 | 2,3020 |
| 10 | 0,425 | 0,639 | 4,2720 | 3,0870 |

**Appendix 4: Overview of Cdc17[1-175]-GFP fusions, expressed from pREP3X, and their localisation in *S. pombe* Sp5 cells**